PERFORMANCE COMPARISON OF MULTISPECTRAL DATA COMPRESSION TECHNIQUES

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

Master of Technology

by

Vineeta Srivastava

to the

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

February 1996

EE-1996-M-SRI-PER

Certificate

It is certified that the work contained in the thesis entitled **PERFORMANCE COMPARISON OF MULTISPECTRAL DATA COMPRESSION TECHNIQUES**, by Vineeta Srivastava , has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

February 1996

Dr. Sumana Gupta

(Assoc) Professor

Department of Electrical Engineering

I.I.T. Kanpur

Abstract

It is well known that present day remote sensing satellite systems are often constrained by downlink communication bandwidth. To meet this challenge collected image data need to be compressed efficiently. As the collected images exhibit a high degree of spatial and spectral redundancy , in the present work exploitation of these redundancies are utilized to achieve high compression.

A 3-D transform based compression technique is developed and its performance is compared with other reported algorithms. [28] [1]

The compression system used comprises of two subsystems : a spectral dcorrelation subsystem followed by a spatial decorrelation subsystem. The spectral decorrelation is carried out in three stages (a). data partitioning for terrain adaptive approach only (b). The optimum KL transform for energy compaction of data and (c). mapping of the eigen planes to eigen images . The decorrelated eigen images are fed to the spatial decorrelation module which consists of embedded zerotree coder based on wavelet transform. Wavelet transform exploits human visual perception deficiencies and offers the opportunity for high compression ratios.

Compression system has been evaluated on two typical test image sets having 4 bands ranging from 0.45μm to 0.8μm and 7 bpp. These test images contain a large variety of natural and urban terrains. Spectral decorrelation subsystem results in two significant eigenbands out of four i.e. two eigen bands have negligible variance or information content. Spectral fidelity of the images is also preserved after reconstruction.

The maximum loss in correlation coefficient matrix is only 7 percent. These eigen bands are furthur spatially decorrelated using embedded zerotree coding method. The reconstructed image fidelities ranges from near lossless at about 10:1 compression ratio i.e. 0.7 bpp to visually lossless upto 80:1 compression ratio i.e. 0.09 bpp . Compared to the compression ratio obtained in [24] based on KLT/DCT methods , the performance of the present technique is far superior.

In this compression system, spectral and spatial modularity allows EZW or KLT to be replaced by any other spatial or spectral coding procedure.

The embedded zerotree coding used , have further advantages. There is no training of any kind and no ensemble statistics of the images are used in any way. An interesting property of embedded coding is that when the encoding and decoding is terminated during middle of a pass or in the middle of the scanning, there are no artifacts produced that would indicate where the termination occur.

To

My Parents

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisor Dr. Mrs.Sumana Gupta for her guidance and encouragement at each step and also for her useful suggestions.

My sincere thanks are also due to Dr. Onkar Dixit and Anand sirohi for providing the multispectral image data. I am also grateful to Dr.P.C.Das with whose help I could get the literature on wavelets.

Lastly I also wish to express my indebtedness to all my faculty members and friends for their co-operation.

Vineeta Srivastava

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Remote sensing systems currently in use and future systems are likely to be constrained in terms of downlink communication bandwidth. To overcome this problem images obtained from satellite and other airborne multispectral collection platforms need to be compressed . Multispectral image (MSI) compression has evolved to meet this challenge.

This multispectral imagery can be used in many applications covering areas in global environmental monitoring , mapping , charting, geodesy and land use planning. There is also the area of natural resourse management including forestry, agriculture, water quality monitoring and wild life habitat management. With such a diverse set of application comes a wide variety of system requirement. These systems require better spatial, spectral and radiometric resolution. The volume of data increases enormously with the above requirements.

1.1 Motivation for compression

Need for various compression schemes arises due to the massive increase anticipated in the volume of the remote sensing data coupled with existing limitations on channel bandwidth. As an example, it is well known that the Earth Observing System carrying a High Resolution Image Spectrometer (HIRIS) with 192 spectral channels generate 100 Mbp/s of data. Similarly the HRMSI (High Resolution Multispectral Imagery) which has 4 multispectral bands and a pan band , with 10 mt GSD and pan band collecting 5 mt. GSD with radiometric precision of 9 bits per pixel for pan band and 10 bits per pixel for the multispectral band and swath width of 40 km. generate 225 Mbps of data. Advanced system carrying 18 spectral bands , 5 mt.GSD and 100 km. of swath generate 4.28 Gbps of data.

Such dynamic increase in the volume of data transmitted has made it imperative to develop novel compression algorithms which can compress remote sensing imagery data to manageable rates.

1.2 Types of compression schemes

Two forms of data compression are possible. The first is lossless or reversible coding method. This method permits exact reconstruction of the original data from the compressed format. Examples are Huffman codes and run-length code. The reduction in volume of data achievable through this is limited and rarely exceeds three.

The other type of compression technique is the lossy or irreversible compression

method. This method often introduces error or distortion in the reconstructed data. It can provide significantly higher compression ratios ranging from 8 to 30 with very little loss. The goal of lossy data compression is to obtain the best possible reproduction for a given bit rate or equivalently to minimize the rate required for a given fidelity . The error introduced can be minimized so that the reconstructed image is visually indistinguishable from the original.

The images within a multispectral data set often exhibit high correlations between spectral bands in addition to the spatial redundancy present within a band. Exploitation of the redundancies present result in a large compaction of the data.

Recent work in the area of multispectral bandwidth compression can be categorized into three groups.

(1). Vector Quantization based technique : The vector nature of registered MSI makes it naturally favourable for VQ . However one of the problem associated with vector quantization is the high encoding complexity which grows exponentially with increase in vector dimension.

(2). Predictive Techniques : With high band to band redundancy, there is an enhanced capability to predict the pixels in one band, given knowledge of adjacent bands. Three dimensional prediction structures can be constructed by weighting neighbouring pixels from within the same band and from adjacent bands to predict the current pixel value. The residual value between the predicted and actual value is then quantized and compressed.

(3). 3-D transform based technique : In this scheme KL transform technique is generally used to decorrelate the spectral bands. The eigen bands are further decorrelated

spatially by using general 2-D transform coder such as the JPEG. Recently wavelet transform has been used for spatial decorrelation [10]

In general, transform based technique outperforms other prediction and vector quantization based technique. Results reported in [28] shows that at bit rate greater than 1.5 bpp , 3-D transform based technique and other two schemes give the same performance. At rate less than 1.5 bpp , 3-D transform based technique performed better as compared to other methods.

1.3 Objective of the Thesis

The primary objective of the thesis is to develop and test new compression algorithms to obtain the best quality multispectral imagery for a given bit rate.

The first algorithm uses 3-D transformation coding technique in which the KL transform is used to exploit the spectral correlation followed by wavelet transform with embedded zerotree coding of its coefficients to exploit the spatial redundancy of the multispectral image set.

In the second algorithm the spatial decorrelation is carried out by using wavelet packet transform with embedded zerotree coding.

1.4 Organization of the thesis

The thesis has been presented in six chapters. In chapter I we introduce the remote sensing system, requirement of multispectral imagery and need for compression. We briefly discuss the conventional methods for multispectral image compression.

Chapter II discusses the overall compression system, encoder and decoder. Chapter III is concerned with methods of spectral decorrelation.

In chapter IV a brief review of the wavelet transform and wavelet packet transform is given. The method to select the best basis coefficients of wavelet packet library is discussed. Also the method of embedded zerotree coding using above transforms is discussed. Lastly adaptive arithmetic coding is reviewed , which is being used to code the zerotree structure.

Chapter V gives the simulation results of the algorithms developed in previous chapters and compares the performance of these algorithms with other compression schemes.

Chapter VI concludes the method adopted and the results achieved in this thesis.

Chapter 2

Multispectral image compression system : An Overview

A simplified version of most compression system can be easily described by a conceptual flow through three different stages, transform ,quantization and compression. fig 2.1 shows it. This concept is useful in describing the functional parts of a compression algorithm.

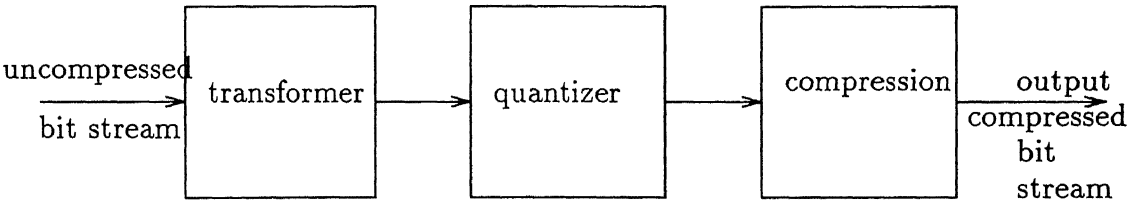


Figure 2.1: Basic compression system

The first step : a transform represents a method to convert image pixels into a form that is more amenable to quantization. This transform can take on many different forms from linear prediction or orthogonal transform to vision modelling. This is typically considered a " decorrelation" step because in the case of unitary transformation the resulting transform coefficients are relatively uncorrelated. The primary purpose of this is to allow each of the coefficients to be quantized independently of the others.

Most unitary transform also have the tendency to pack a large fraction of the average energy of the image into relatively few components of the transform image. Since the total energy is preserved , this means many of the transform coefficients will contain very little energy and can be highly compressed.

The transformed pixel values are real numbers which must also be approximated in a finite alphabet, or quantized before they can be transmitted. All the distortion from a lossy transform coding scheme is introduced at this step. The range of transformed value is divided up into numbered subinterval or bins. Any pixel value falling into a bin is approximated by bin's index.

The last step of the process , the coding step is where the actual compression takes place . Having gone through an entropy reducing step, the resultant quantized values can be coded to reduce the overall symbol length.

2.1 MSI properties relevant to compression

Though the nature of the distortion introduced into the data varies with the type of algorithm, all data compression schemes operate by exploiting redundancy inherent to data.

2.1.1 Spatial Properties

Pixel to pixel correlation is the form of spatial redundancy that is most often exploited by single band image compression algorithms. The normalized autocovariance function for pixels $I(m,n)$ in an image is defined by

$$C(m,n)^2 = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \frac{I(k,l)I(k+m,l+n) - {\mu_I}^2}{{\sigma_I}^2} \tag{2.1}$$

where

$$\mu_I = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m,n) \tag{2.2}$$

$$\sigma_I^2 = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I^2(m,n) - \mu_I^2 \tag{2.3}$$

The behaviour of autocovariance function shows the redundancy present in data.

2.1.2 Spectral Properties

MSI compression techiniques can be designed to exploit redundancies that exist between spectral bands. Provided that the spectral bands are completely registered, significant band to band correlation is present even across large difference in wavelength.

2.2 Overview of the compression system

Here 3 dimension transform based scheme is used to compress the multispectral imagery.

2.2.1 Encoder

The encoder block diagram is shown in fig 2.2.

The compression system is having two modules. One is spectral decorrelation module and other is spatial decorrelation module. This module based compression system scheme allows to replace any module by any different coder.

(a) Spectral decorrelation module

It consists of three parts.

(1). Data partitioning module (optional) :

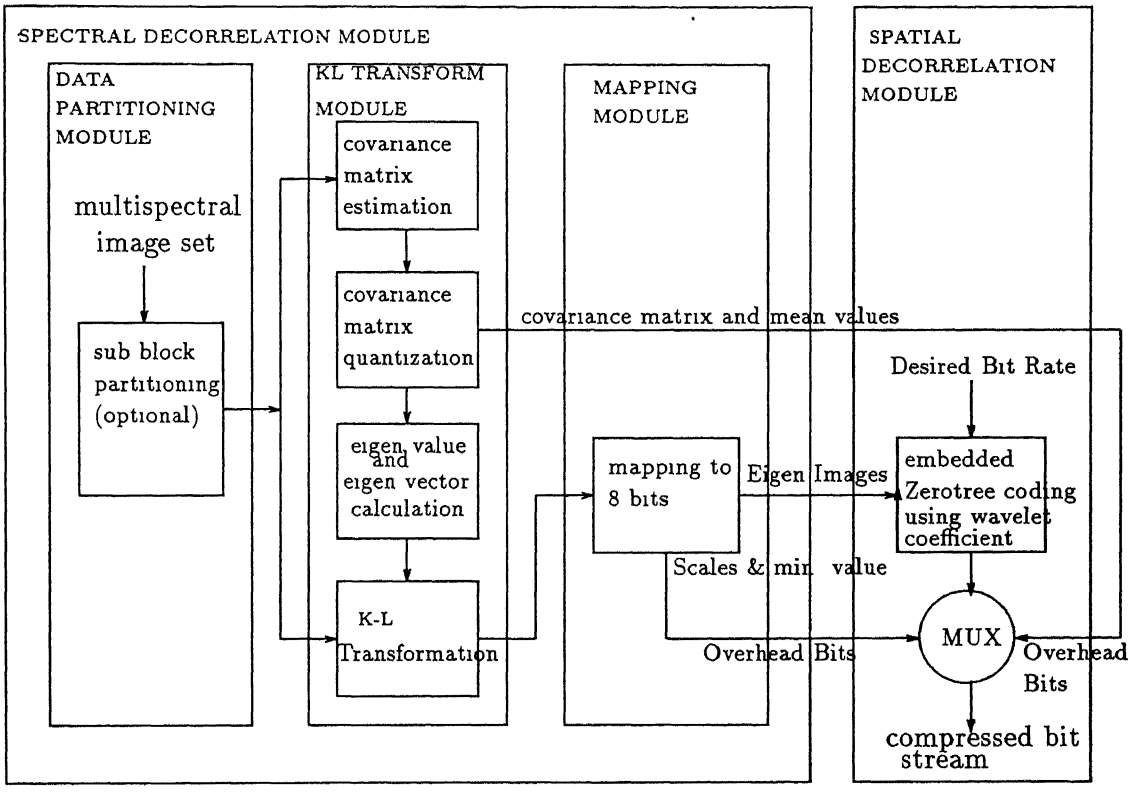


Figure 2.2: Encoder block diagram for multispctral image compre ssion

In the data partitioning module the set of multispectral images are partitioned into sets of non-overlapping images : sub-block sets , which are sequentially fed to the KL transformation module for spectral decorrelation. This module is necessary for terrain adaptive compression system but not required for non-terrain adaptive compression system.

(2). KL transformation module :

In the KL transformation module the multispectral sub-block set is spectrally decorrelated to produce a set of eigen planes. The basis function for the KLT are the eigen-vectors of the autocovariance matrix associated with multispectral sub-block set. The covariance matrix is estimated first and then quantized to two bytes per element. The eigen planes are formed by matrix multiplication of multispectral block and basis function (subblocks and associated basis functions for terrain adaptive case)

(3). Mapping module :

The eigen planes are in floating point format and assume both positive and negative values. In this module eigen plane set is converted into the 8 bits eigen image set via linear/non-linear mapping of each plane into the 0-255 range. These spectrally decorrelated eigen images are fed to spatial decorrelation module.

(b). Spatial decorrelation module

To spatially decorrelate the eigen image set, the eigen images are sequentially embedded coded using zerotrees of wavelet coefficients. The wavelet transform is used to exploit the spatial redundancy . The main contribution of the wavelet theory and multiresolution analysis is that it provides an elegant framework in which both the signals, which are either localized in frequency or space can be analyzed with equal resolution. This wavelet transformed image is significant map encoded by predicting the absence of significant information across scales by exploiting self similarity inherent in the images. This uses a new data structure called zerotree. This significance map is quantized with successive approximation quantization and further coded by lossless data compression scheme which is achieved via adaptive arithmetic coding.

2.2.2 Decoder

The block diagram of decoder is shown in fig 2.3. The mapping information and covariance matrix are extracted from the received bit stream. The eigen images are reconstructed from their compressed bit stream using EZW decoder. The mapping information is used to span the 8 bit dynamic range of the decoded eigen images to their original range. The inverse KLT basis functions are obtained from the extracted covariance matrix via transposing its eigen-vectors. An inverse KLT is performed to produce the reconstructed multispectral image set. In terrain adaptive case sub-blocks

are mosaiced together to reconstruct the original image set.

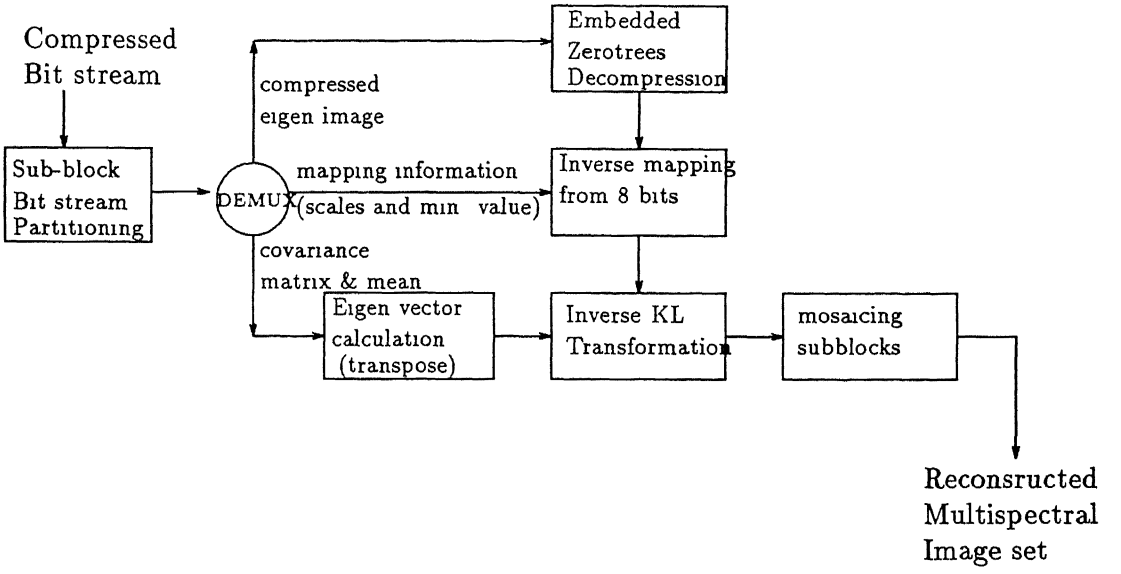


Figure 2.3· Block diagram of decoder

Chapter 3

Spectral decorrelation subsystem

Removing inherent spectral correlation in the data results in a significant compaction of data to be coded.

3.1 Spectral decorrelation by KLT

Figure 3.1 illustrate the KLT orthogonal transformation. KLT is the optimum method to decorrelate but it does not have any fixed basis. KLT is reviewed below.

For a real Nx1 random vector x, the covariance matrix of the x vectors is defined as

$$C_x = E[(x - m_x)(x - m_x)'] \tag{3.1}$$

where

$$m_x = E[x] \tag{3.2}$$

is the mean vector, E is the expectation operator and the prime (') indicates transposition. Equations 3.1 & 3.2 can be approximated from the samples by using the relations.

$$m_x \cong \frac{1}{M} \sum_{i=1}^{M} x_i \tag{3.3}$$

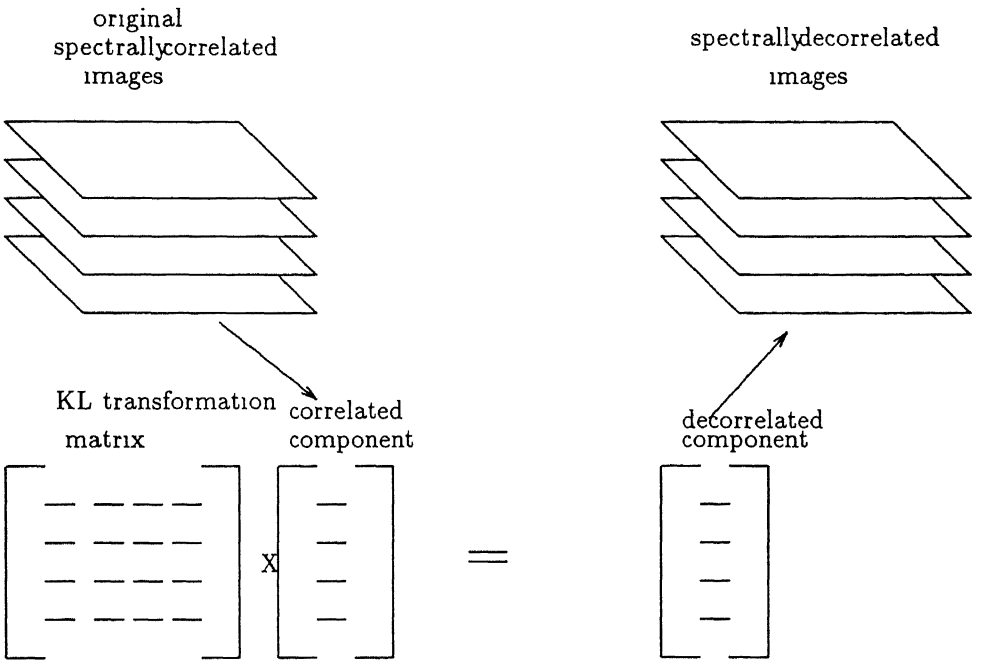


Figure 3.1: Removing spectral correlation via KL transformation

and

$$C_x \cong \frac{1}{M}\sum_{i=1}^{M}(x_i - m_x)(x_i - m_x)' \tag{3.4}$$

or equivalently

$$C_x \cong \frac{1}{M}[\sum_{i=1}^{M} x_i x_i'] - m_x m_x' \tag{3.5}$$

let e_i and λ_i, i=1,2,.....N be the eigen vectors and corresponding eigenvalues of C_x. It is assumed for convenience in notation that the eigen values have been arranged in decreasing order so that $\lambda_1 \geq \lambda 2 \geq \lambda_3 ..... \geq \lambda_N$. A transformation matrix whose rows are the eigen vectors of C_x is given by

$$A = \begin{bmatrix} e_{11} & e_{12} & . & . & e_{1N} \\ e_{21} & e_{22} & . & . & e_{2N} \\ . & . & . & . & . \\ e_{N1} & e_{N2} & . & . & . \end{bmatrix} \tag{3.6}$$

where e_{ij} is the j^{th} component of the i^{th} eigen vector. The KL transform then consists

simply of multiplying a centralized image vector $(x - m_x)$ by A to obtain a new image vector y; i.e.

$$y = A(x - m_x) \tag{3.7}$$

Mean of the transformed vector y is equal to zero vector as can be shown directly

$$\begin{aligned} m_y &= E[y] \\ &= E[A(x - m_x)] \\ &= AE[x] - Am_x \\ &= 0 \end{aligned}$$

covariance matrix of y, C_y can be expressed in terms of C_x as

$$\begin{aligned} C_y &= E[(Ax - Am_x)(Ax - Am_x)^{'}] \\ &= E[A(x - m_x)(x - m_x)^{'}A^{'}] \\ &= AE[(x - m_x)(x - m_x)^{'}]A^{'} \\ &= AC_xA^{'} \\ &= \Lambda \\ &= \text{diagonal} \end{aligned}$$

C_y is a diagonal matrix which shows that all the coefficient representing inter-band correlation are zero or the KL transform coefficient are completely uncorrelated ., so this is called optimal transform.

For the N spectral images to be decorrelated, we form an N by N KL transformation matrix composed of the N ordered eigenvectors of the associated autocovariance matrix . Each vector of the spectrally-correlated components form an output vector of spectrally decorrelated components . The output vectors are placed adjacent to one

another, in the same order as the input vectors, to form the stack of the spectrally decorrelated eigen planes.

3.2 Quantization of the eigen planes

the spectrally decorrelated eigen planes are represented in floating point real numbers. They must be quantized to 8 bit, 10 bit or 12 bit to be subsequentially.

3.2.1 Linear quantization for multispectral imagery

For multispectral imagery all eigen planes are linearly quantized into 8 bit images. They are the eigen images . Alternatively high dynamic range, low order eigen planes may be quantized into 10 bits or 12 bits but practically no appreciable gain in performance is achieved. For the 10/8 bit quantizer , only the first two eigen planes are quantized with 10 bit quantizer.This gives 5% increase in bitrate and reduces mean square error from 0.3 to 0.28 compared to nominal 8 bit quantization of the eigen image.

It can be concluded that dual 10/8 bit quantizer does not offer better results because (i) the maximum induced error is only 1 count in 256. for both the cases (ii) the gain in quantization round-off is smaller than the error introduced at the next stage (iii) it requires more bit rate.

3.2.2 Nonlinear/Linear quantization for high dynamic range

For high dynamic range (12 bit) imagery, an optimal 8 bit or 12 bit quantizer may be used to quantized the first one or two high dynamic range eigen planes. Lower eigen planes may be quantized using a linear 8 bit quantizer. The high dynamic range of the first few eigen planes typically have an uneven (unequalized) histogram with long

narrow tails. A nonlinear quantizer achieves a substantially lower overall quantization error by allocating finer quantization steps to the large central portion of the dynamic range.

3.3 Terrain adaptive approach

A typical multıspectral image set obtained from either satellite or airborne collection platforms exhibits a number of different terrains such as water, forest, cloud etc. Each terrain has unique spectral signature. Thus to achieve the highest compactness of spectral information, the spectral transformation parameters must adapt to the local terrain characteristics. In the terrain adaptive approach the covariance matrix, from which the spectral transformation basis functions are derived, is updated frequently. The smaller the block size over which the covariance is updated is , the more efficient is the spectral decorrelation process. The drawback with the selection of a small sub-block is the resulting increase in overhead bit rate due to an increase in the number of sub-blocks and it is more computationally intensive . Resulting eigen images clearly indicates the blocking effect. Thıs approach also results in decrease in variance. This decrease in variance shows much greater compaction of data and substantially lower performance.

3.4 Overhead Information

The overhead information at spectral decorrelation module contains the mean and covariance matrix and eigen plane quantization parameters. For non-terrain adaptive (no sub-blocking of original data) spectral decorrelation of N band multispectral image, overhead is as given below.

(a). Mean : m_x requires 2 bytes per band i.e. total bit required is 16N.

(b). Covariance matrix is diagonally symmetric only non-redundant half is transmitted. Size of covariance matrix is NxN and the no. of non-redundant element is $[N^2/2 + N/2]$ the total bit requirement is $8[N^2 + N]$, when 16 bits per minimum value.

(c). Quantization parameters consists of scale factor and minimum value. It requires 16N bits ,if integer value is transmitted for minimum value and 8 bits for scale factor. Note that scale of unity is not transmitted.

Total bit requirement is

$$8 + 16N + 8(N^2 + N) + 16N \tag{3.8}$$

Overhead bit rate B_{oh} in bpp is

$$b_{oh} = \frac{8 + 40N + 8N^2}{NB^2} \tag{3.9}$$

where the size of the image is BxB.

For terrain adaptive approach

The overhead is K^2[Nonterrain adaptive approach] where K is $\frac{dim-length}{\text{window size}}$ or K^2 is no. of subblock present.

3.5 Alternative decorrelating schemes for spectral decorrelation

KL transform has certain disadvantages as

(1). Computation burden is high as this transform is not having any fixed basis, every time covariance matrix has to be estimated and eigen vectors has to be calculated. Furthermore no fast algorithm exists for it.

(2). Overhead is also great in this case as whole basis information has to be transmitted along the basis.

Because of these disadvantages alternative spectral decorrelation techniques are also tried.

3.5.1 Discrete Cosine Transform

Discrete cosine transform (DCT) can be used instead of KLT. The advantage with DCT is that the basis function is fixed regardless of the characteristics of the data. Unlike the KLT , it does not require covariance and eigen vector caculation and , as such is much easier to represent. The drawback with this scheme is that the spectral decorrelation efficiency will be significantly poor. KLT approach including overhead will result into lower bit requirement.

3.5.2 Affine prediction

This approach is based on prediction in spectral band and this is simpler to set-up at satellite.Here we review the concept of affine predictor

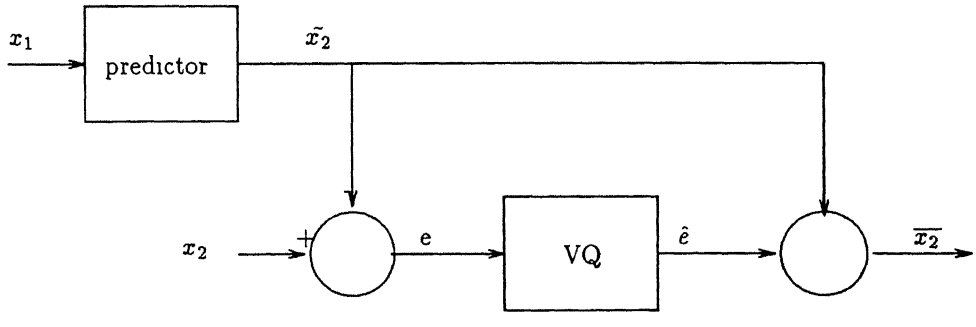


Figure 3.2: Affine predictor for two bands

Considering first the case of two bands labelled band 1 and band 2. The vector is formed $x = (x_1, x_2)$ where x_1 is element from first band and x_2 is the element from the

second band at identical location.

The affine predictor is obtained by restricting the predictor f(.) of x_2 to be of the form

$$\tilde{x}_2 = f(x_1) = Ax_1 + B \tag{3.10}$$

where A and B are scaler quantities.If B is set to zero then affine predictor reduces to the well known linear predictor , the best affine predictor employs A and B chosen to minimize the distortion over all possible A and B . The solution to the affine prediction problem is derived readily as a simple extension of the linear predictor solution and given by

$$A = \frac{E[(x_2 - E(x_2))(x_1 - E(x_1))]}{C_{x_1}} \tag{3.11}$$

where

$$C_{x_1} = E[(x_1 - E(x_1))(x_1 - E(x_1))] \tag{3.12}$$

and

$$B = E[x_1] - A[E[x_1] \tag{3.13}$$

where E is expectation operator.

Affine predictor for the two band can be extended to more than two bands.For J bands consider $(x_1, x_2, x_3, ...x_J)$ as one vector. Here x_i is the element from band i. Feature band to be quantized is selected as a subset of these J bands. The remaining bands are estimated from quantized feature band.

In the simplest case only one of the band i is chosen as the feature band . Rest each of the bands is predicted separately from the quantized feature band. For the J band multispectral data set it requires the design of J-1 error vectors. This predictor has advantage of simplicity in implementation in satellite but this predicts one band from only single band. As in multispectral imagery bands are correlated with many other

bands , this predictor does not exploits the redundancy present in spectral domain. The performance of this scheme is significantly lower than KLT.

Chapter 4

Spatial Decorrelation Subsystem

Spatial redundancy is inherent in each of the bands of the multispectral imagery. Hence it is necessary to exploit this redundancy in order to achieve significant compression of multispectral data.

In image coding traditional transform coders such as those using the DCT, decompose the images into a representation in which each coefficient corresponds to a fixed frequency bandwidth where the bandwidth and spatial area are effectively the same for all coefficients in the representation. Edge information tends to disperse so that many non-zero coefficients are required to represent the edge with good fidelity. However , since the edges represent relatively insignificant energy with respect to entire image , these transform coders have been fairly successful at medium and high bit rates. At extremely low bit rates these coders tend to allocate too many bits to the "trends" or the signal behaviour that is more localized in frequency but persists over a large number of lags in time domain and few bits are left for "anamolies" or the signal behaviour that is more localized in time or space domains and tends to be wide band in frequency domain. As a result blocking artifacts often result. Wavelet transform technique shows promise at extremely low bit rates.

In this chapter we review the basic concept of time frequency analysis , wavelet transform followed by wavelet packet transform. We discuss the wavelet transform technique followed by the embedded zerotree coding method to decorrelate spatially the eigen images obtained from the spectral decorrelation subsystem.

4.1 Wavelet transform : A review

4.1.1 Continuous wavelet transform

The continuous wavelet transform of signal s(t) by definition is convolution with a wavelet $\psi(t)$ dilated by a factor a

$$W_s(a,b) = a^{-1/2} \int s(t)\psi(\frac{t-b}{a})dt \tag{4.1}$$

$$W_s(a,b) = a^{-1/2} \int S(w)\psi(aw)e^{jwb}dw \tag{4.2}$$

which is equivalent to filtering the signal s(t) with bandpass filter $\psi(aw)$, whose bandwidth changes according to scale parameter a, clearly, large scale corresponds to narrow smoothing filters that represented a global view of the signal s(t) and small scales correspond to filters that look into the details of s(t) (i.e. high frequency components).

Wavelet expansion of the signal s(t) is essentially a decomposition of its frequency content using filters of constant relative bandwidth. The signal can be recovered from its wavelet coefficients using

$$s(t) = a^{-5/2} \int_a \int_b W_s(a,b)\psi(\frac{t-b}{a})dadb \tag{4.3}$$

assuming that

$$\int_t \psi(t) = 0 \tag{4.4}$$

$$\int_0^\infty |\frac{\psi(w)^2}{w}| dw < \infty \tag{4.5}$$

As it is the case with the fourier transform where a signal is expanded in terms of complex exponentials of different frequencies. A wavelet transform involves dilation of a single (mother) wavelet .The choice of a mother wavelet depends on the application, where a particular wavelet is chosen based on its time and frequency localization.

Orthogonality is an important element of wavelet analysis and a mother wavelet is orthogonal to its own dilations and translations. Wavelets provide orthonormal basis for expansion of functions that are not of single frequency and are therefore ideal for characterizing signals with discontinuities.

4.1.2 Properties of wavelets

1. $\psi(w)$=0 at w=0, or equivalently $\int \psi(t)dt = 0$ i.e. they have zero d.c. components.
2. They are band pass signals.
3. They decay rapidly towards zero with time.

Property (1) is a consequence of admissibility condition of the wavelets, the condition that ensures the wavelet transform has an inverse. The rapid decay of $\psi(t)$ is not necessary theoretically to be a wavelet. However $\psi(t)$ in practice should have compact support in order to have good time localization.

4.1.3 Discrete Wavelet transform

The wavelet transform parameters can be discretized so that

$$C_{m,n} = a_0^{-m/2} \int_i s(t)\psi(\frac{t - na_0^m T}{a_0^m}) \tag{4.6}$$

where

$$a = a_0^m \tag{4.7}$$

$$b = na_0^m T \tag{4.8}$$

and T is sampling period. The signal s(t) can be recovered from its expansion coefficients using

$$s(t) = A \sum_m \sum_n C_{m,n}\psi(\frac{t - na_0^m T}{a_0^m}) \tag{4.9}$$

where A is constant.

The case where $a_0 = 2$ is known as the dyadic wavelet transform, where the signal s(t) is bandpass filtered using octave band filters. This type of wavelet has the form

$$\psi_{m,n}(k) = 2^{-m/2}\psi(2^{-m}k - n) \qquad m, n \subset Z \tag{4.10}$$

The discrete wavelet transform of discrete time sequence s(k) is essentially a multiresolution characterization of the s(k). Generally we take the discrete wavelet transform of a signal that is both time limited and frequency limited. A continuous time signal, uniformly sampled satisfies this criterion. A dyadic discrete wavelet transform is essentially a decomposition of the spectrum of s(k); S(w) into orthogonal subbands defined by

$$\frac{1}{2^j T} \leq w \leq \frac{1}{2^{j+1} T} \tag{4.11}$$

where j=1,2,....J and T is sampling period associated with s(k).

4.1.4 Time–frquency representation

In the time-frequency analysis of a non-stationary signal, there are two conflicting requirements. That is the window width must be long enough to give the desired frequency resolution but must be short enough so as not to blur the time dependent events.

Define the bandwidth Δf of the filters as

$$\Delta f^2 = \frac{\int f^2 |G(f)|^2 df}{\int |G(f)|^2 df} \tag{4.12}$$

and spread in time is given by

$$\Delta t^2 = \frac{\int t^2 |g(t)|^2 dt}{\int |g(t)|^2 dt} \tag{4.13}$$

Now resolution in time and frequency cannot be arbitrary small, because their product is lower bounded by

$$\Delta t \Delta f \geq \frac{1}{4\pi} \tag{4.14}$$

Short time fourier transform (STFT) uses fixed window width. So this represents fixed time and frequency resoluton. However by varying the window used one can trade resolution in time for resolution in frequency. Time- frequency diagram shows the effect of shift in time τ and modulation by w_0 in frequency. Scaling the function by a or $f'(t)=f(at)$ results in $I'_t=\frac{1}{a}I_t$ and $I'_w=aI_w$, that is both shape and localization of the tile have been affected.

For Wavelet transform for the large a basis function $\psi_{ab}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})$ becomes a stretched version of the prototype wavelet that is a low frequency function, while for small a, the basis function becomes a contracted wavelet that is a short high frequency

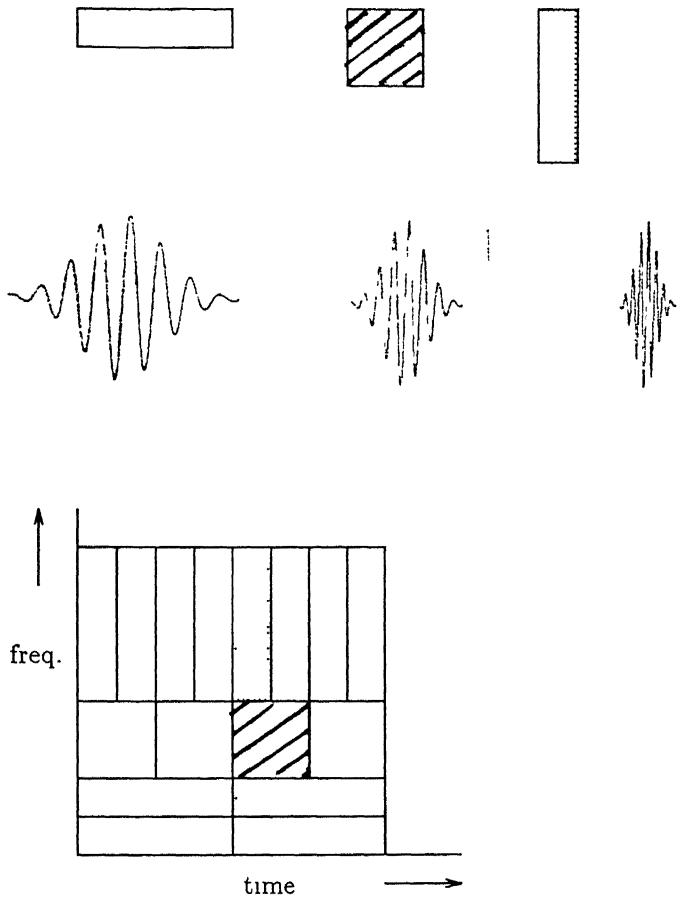


Figure 4.1: Time-frequency representation of wavelet transform

function. The frequency resolution of the wavelet transform involves a different trade-off : at high frequency, the wavelet transform is sharper in frequency. Time frequency diagram of wavelet transform is shown in fig 4.1.

4.1.5 Multiresolution Analysis

The idea of multiresolution analysis (MRA) is very similar to subband decomposition and coding. In the dyadic case the wavelet transform is actually an octave band filter, can be interpreted as a constant Q filtering with a set of octave band filters followed by sampling at the respective Nyquist frequencies.

Multiresolution analysis is the decomposition of a signal into components of different scale (frequency) of 2^{-m}, m integer. Associating with each scale (frequency band) is a subspace V_m. These subspaces are time functions satisfying

1. Containment

.....$V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2}$..... The subspace begin with null space and expands in scale of two to reach the space of all square integrable function. In particular if $x(t) \in V_{-1}$ then $x(2t) \in V_{-1}$.

2. Existence of orthonormal scaling functions

There exists a scaling function $\phi(t) \in V_0$ s.t. $\phi_{m,n}(t) = 2^{-m/2}\phi(2^{-m}t - n)$, n integer forms an orthonormal basis that spans V_m. i.e. V_0 is the space of all band limited functions with frequency interval $(-\pi, \pi)$, set of function $\phi(t - n)$ forms an orthonormal basis for V_0. Similarly V_{-1} is the space of band limited functions with frequency in the interval $(-2\pi, 2\pi)$. Here $\phi(t)$ is called the scaling function because it derives an approximation in V_0 of signals in V_{-1}. The scaling function $\phi_{m,n}(t)$ forms an orthonormal basis in V_m can be exactly represented as a linear combination of $\phi_{m,n}(t)$.

3. Basis function defined by two scale difference equation:

Since $\phi_{0n}(t)$ spans V_0 and $\phi_{-1n}(t)$ spans V_{-1} and V_{-1} contains V_0 , $\phi_{0,0} = \phi(t)$ is a linear combination of $\phi_{-1,n}(t) = \sqrt{2}\phi(2t - n)$

i.e.

$$\phi(t) = 2\sum_{l=0}^{p-1} g(l)\phi(2t - l) \tag{4.15}$$

which is a two scale difference equation. Now call W_0 the space of bandpass functions

with frequencies in the interval $(-2\pi, \pi) \cup (\pi, 2\pi)$ then

$$V_{-1} = V_0 \oplus W_0 \tag{4.16}$$

that W_0 is the orthogonal complement in V_{-1} of V_0. In other words V_{-1} is equivalent to V_0 plus some added detail corresponding to W_0.

Generalizing V_{-i} is the space of band limiting functions with frequency in the interval $(-2^{-i}\pi, 2^{i}\pi)$ then we get the relations

$$V_i \subset V_{i-1} \qquad\qquad i \in Z \tag{4.17}$$

and

$$V_{i-1} = V_i \oplus W_i \tag{4.18}$$

where W_i is the space of bandpass functions with frequencies in the interval $(-2^{-(i+1)}\pi, -2^{-i}\pi) \cup (2^{-i}\pi, 2^{-(i+1)}\pi)$.

iterating,

$$V_i = W_{i+1} \oplus W_{i+2} \oplus W_{i+3} \oplus \ldots\ldots\ldots \tag{4.19}$$

finally the direct sum of all W_j's j=i+1,i+2,.....∞ is equivalent to the space of square integrable function band limited to $(-2^{-(i+1)}\pi, 0) \cup (0, -2^{-(i+1)}\pi)$.

It can be seen that V_0 can be decomposed in the following manner

$$V_0 = W_1 \oplus W_2 \oplus \ldots\ldots\ldots W_{j-1} \oplus W_j \oplus V_j \tag{4.20}$$

as can be seen in fig 4.2.

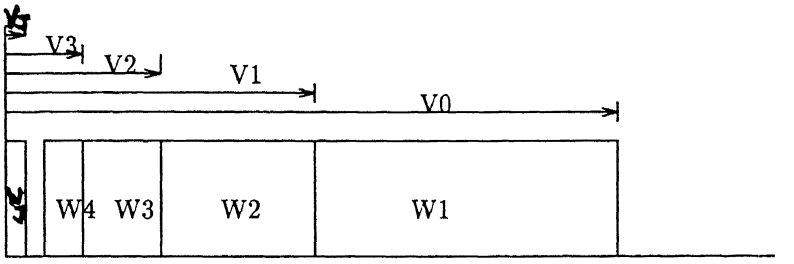


Figure 4.2: Subspace decomposition of square integrable function

Now since $W_0 \subset V_{-1}$, then a function $\psi(t)$ called the wavelet function, whose translates $\psi(t-n)$ spans W_0 can also be written as a linear combination of $\phi_{-1,n}(t)$ which spans V_{-1} thus

$$\psi(t) = 2\sum h(l)\phi(2t-l) \tag{4.21}$$

from the inner product preserving property

$$< \psi_{m,n}(t)\psi_{j,k}(t) >= \delta_{m-j}\delta_{n-k} \tag{4.22}$$

These $\psi_{m,n}(t)$ generate orthogonal subspaces W_m.

(4).Existence of orthonormal wavelet function :

Wavelet function $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$ are orthonormal with

$$< \psi_{m,n}(t)\psi_{j,k}(t) >= \delta_{m-j}\delta_{n-k} \tag{4.23}$$

These $\psi_{m,n}(t)$ generate orthogonal subspaces W_m.

4.1.6 Wavelets and filterbanks

Reffering to the two scale equations 4.15 & 4.21 it is known that the coefficient g(l) and h(l) corresponds to lowpass and highpass filters, as two scale equation are the

interpolation equation by $\phi(2x)$ of the perfect half band lowpass filter. Actually if the filter g(l) is exactly lowpass then h(l) is exactly highpass filter.

The decomposition of V_{-1} into W_0, W_1, W_2, W_3, etc. is essentially a wavelet transform, since it splits one of the resulting half spaces in two etc. Thus if V_{-1} is the space of functions, band limited to $(-2\pi, 2\pi)$ then V_0 and W_0 are the spaces of functions bandlimited to $(-\pi, \pi)$ and $(-2\pi, -\pi) \cup (\pi, 2\pi)$ respectively. That is by iteration the discrete system in fig 4.3 computes exactly the discrete wavelet transform into octave band.

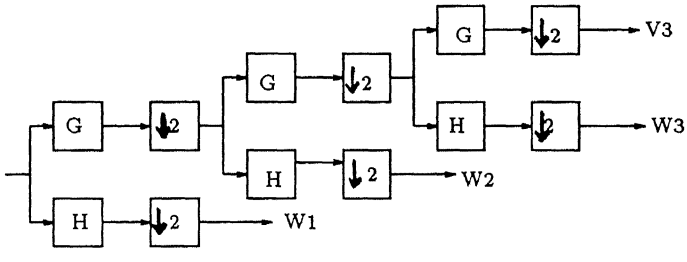


Figure 4.3: Octave band analysis filter bank

Octave band filter bank computes the inner products with the basis functions for $W_1, W_2 \ldots\ldots W_j$ and V_j i.e. we get the orthogonal projection of the input signal onto $W_1, W_2 \ldots\ldots W_j$ and V_j. That is, the input is decomposed into a very coarse resolution (which exists in V_j) and added details (which exists in $W_i, i = 1, 2 \ldots j$) We will call V_j's approximation spaces and W_j's detail spaces. Then the process of building up the signal is intuitively very clear - one starts with its lower resolution version belonging to V_j and adds up the details until the final resolution is reached.

4.1.7 Two dimensional wavelets

The idea is to form a 1-D sequence from the two dimensional image row sequence, do a one dimensional MRA, restore the MRA outputs to a 2-D format and repeat another MRA to the 1-D column sequences. The two steps of restoring to a 2-D sequence and forming a 1-D column sequence can be combined efficiently by appropriately selecting the proper points directly from the 1-D MRA outputs. As seen from fig 4.4 after the 1-D row MRA each low pass and high pass output goes through the 2-D restoration and 1-D column formation process and then move on to another MRA. Let t_1, t_2 be the 2-D co-ordinates and L=lowpass, H=highpass . Then the 2-D separable scaling function is

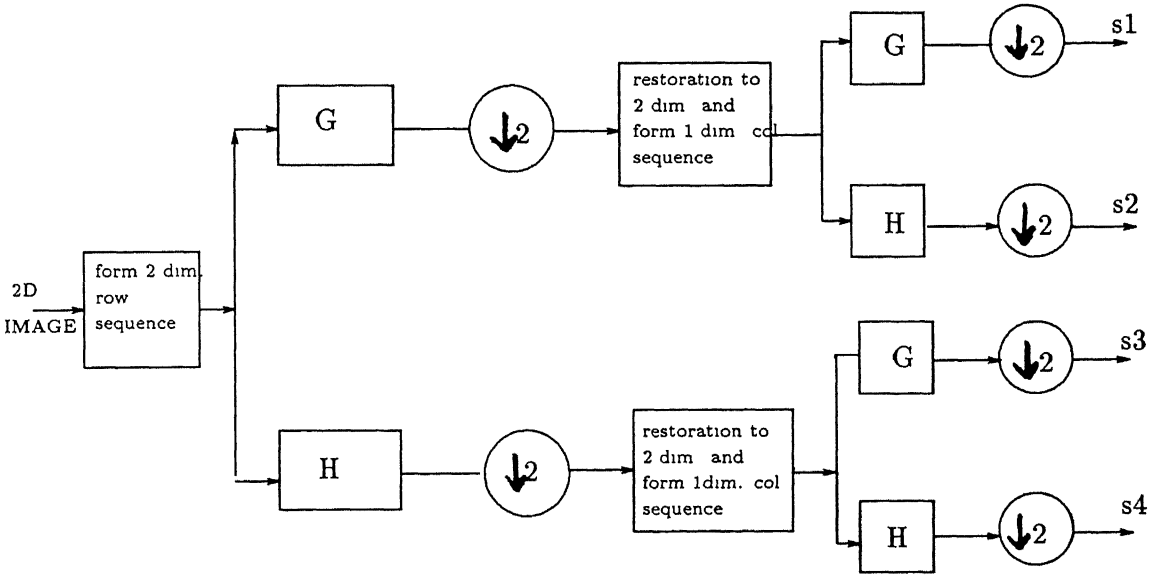


Figure 4.4: 2-D Multiresolution analysis

$$\phi^{(1)}(t_1, t_2) = \phi(t_1)\phi(t_2) \qquad LL \qquad (4.24)$$

and the two dimensional separable wavelets are

$$\psi^{(2)}(t_1, t_2) = \phi(t_1)\psi(t_2) \qquad LH \qquad (4.25)$$

$$\psi^{(3)}(t_1, t_2) = \psi(t_1)\phi(t_2) \qquad HL \tag{4.26}$$

$$\psi^{(4)}(t_1, t_2) = \psi(t_1)\psi(t_2) \qquad HH \tag{4.27}$$

4.1.8 Wavelet packets and generalized filter banks

In MRA ,the input spectrum at each stage is always split into two bands at a time; the higher band becomes one of the outputs while lower band is again split into two bands etc. Now each splitting could be into several bands at a time. In addition, there could be a furthur splitting of the higher bands as well. There is a wavelet interpretation to this generalization of MRA and the outputs are called wavelet packets. Discrete wavelet transform providing an octave band , constant Q analysis are rather restrictive in the available frequency resolution. Wavelet packets, through arbitrary band splitting can have frequency resolutions different from the octave band constant Q scheme. An adaptive system can then choose the most suitable resolutions (combination of wavelet packets) to represent a particular signal.

The defining two-scale equation for wavelet packets are

$$\psi^{(2k)}(t) = 2\sum_{l=0}^{p-1} g(l)\psi^{(k)}(2t - l) \tag{4.28}$$

$$\psi^{(2k+1)}(t) = 2\sum_{l=0}^{p-1} h(l)\psi^{(k)}(2t - l) \tag{4.29}$$

with initial condition

$$\psi^{(1)}(t) = \psi(t) \tag{4.30}$$

where g(l) and h(l) are the same filter coefficients as those for MRA. Wavelet packets come from a linear combination of a mother wavelet $\psi(t)$ as depicted in fig 4.5. In this block denotes linear combination of input in the form of 4.28 and 4.29. Fourier

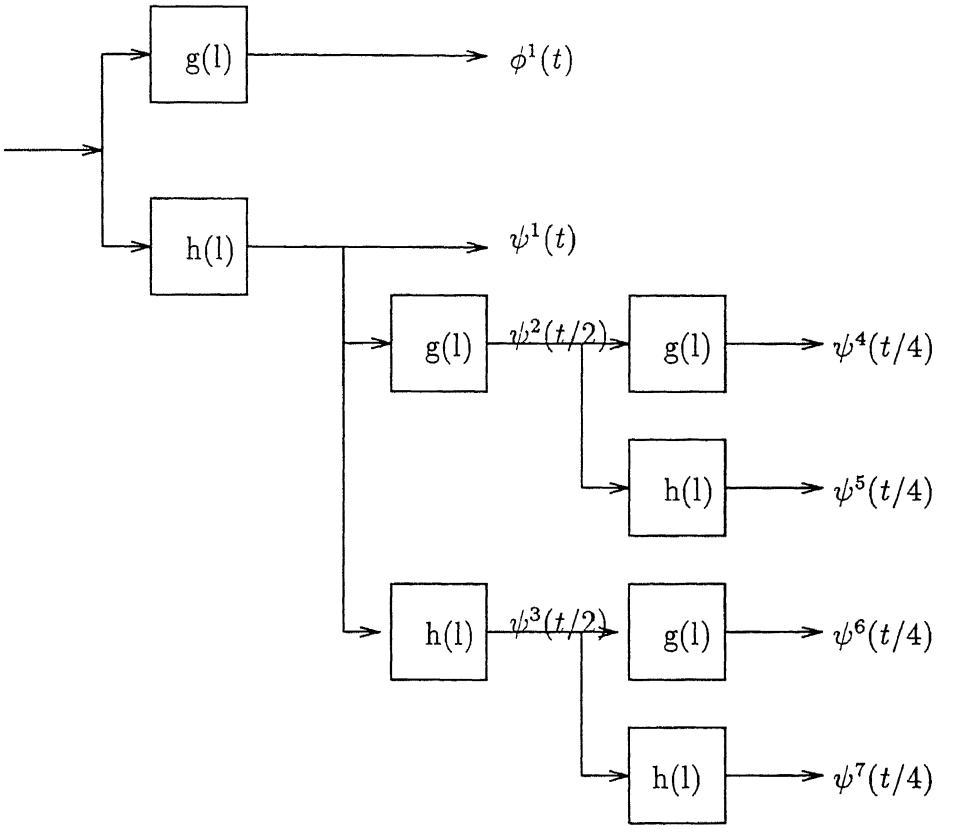


Figure 4.5: Wavelet packet generation

transform of a wavelet packet is

$$\psi^{2^i}(w) = \prod_{l=1}^{i} G(\frac{w}{2^l})\psi(\frac{w}{2^i}) \tag{4.31}$$

$$\psi^{2^i+1}(w) = H(\frac{w}{2})\prod_{l=2}^{i} G(\frac{w}{2^l})\psi(\frac{w}{2^i}), i = 1, 2..... \tag{4.32}$$

other $\psi^{(k)}(w)$ follows similarly.

Octave band filter banks can be generalized to arbitrary tree structures, starting from a single two channel filter banks, all the way from full grown tree of depth J. Figure 4.6 shows all possible tree structure of depth less or equal to two.

In particular full tree, which yields a linear division of the spectrum similar to the short time Fourier transform, and the octave band tree, which performs a two step

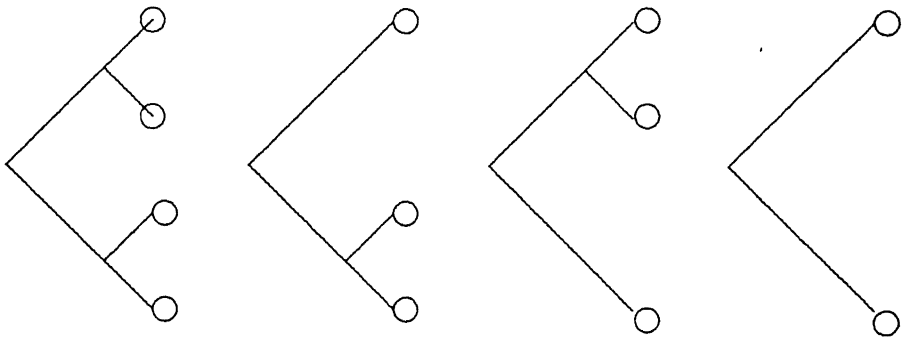


Figure 4.6: Tree structures of depth less than or equal to two

discrete time wavelet series expansion. Such arbitrary tree structure, introduced as a family of orthonormal bases for discrete time signals, known as wavelet packet defined above. The potential of the wavelet packet lies in its capabilities to offer a rich menu of orthonormal bases from which the "best" can be chosen.

Denote the equivalent filters by $g_i^j[n], i = 0, 1, 2....., 2^{j-1}$. In other words $g_i^j[n]$ is the i^{th} equivalent filter going through one of the possible paths of length j.

In the general case, with filter banks of depth J, it can be shown that, counting the no-split tree, the number of orthonormal bases satisfies

$$M_J = M_{J-1}^2 + 1 \tag{4.33}$$

Among this myriad of bases, there are the STFT like bases given by

$$W_0 = g_0^J[n - 2^J k], ..., g_{2^J-1}^{(J)}[n - 2^J k] \qquad k \in Z \tag{4.34}$$

and the wavelet like basis

$$W_1 = g_1^{(1)}[n - 2k], g_1^{(2)}[n - 2^2 k]....., g_1^{(J)}[n - 2^J k], g_0^{(J)}[n - 2^J k] \qquad k \in Z \tag{4.35}$$

It can be shown that the sets of basis funtions in all other bases generalized by the filter bank tree, are orthonormal. Therefore, the impulse responses of all equivalent

filters and their appropriate shifts form an orthonormal bases for $l^2(Z)$.

Time-frequency analysis performed by various filter banks are shown in fig 4.7.

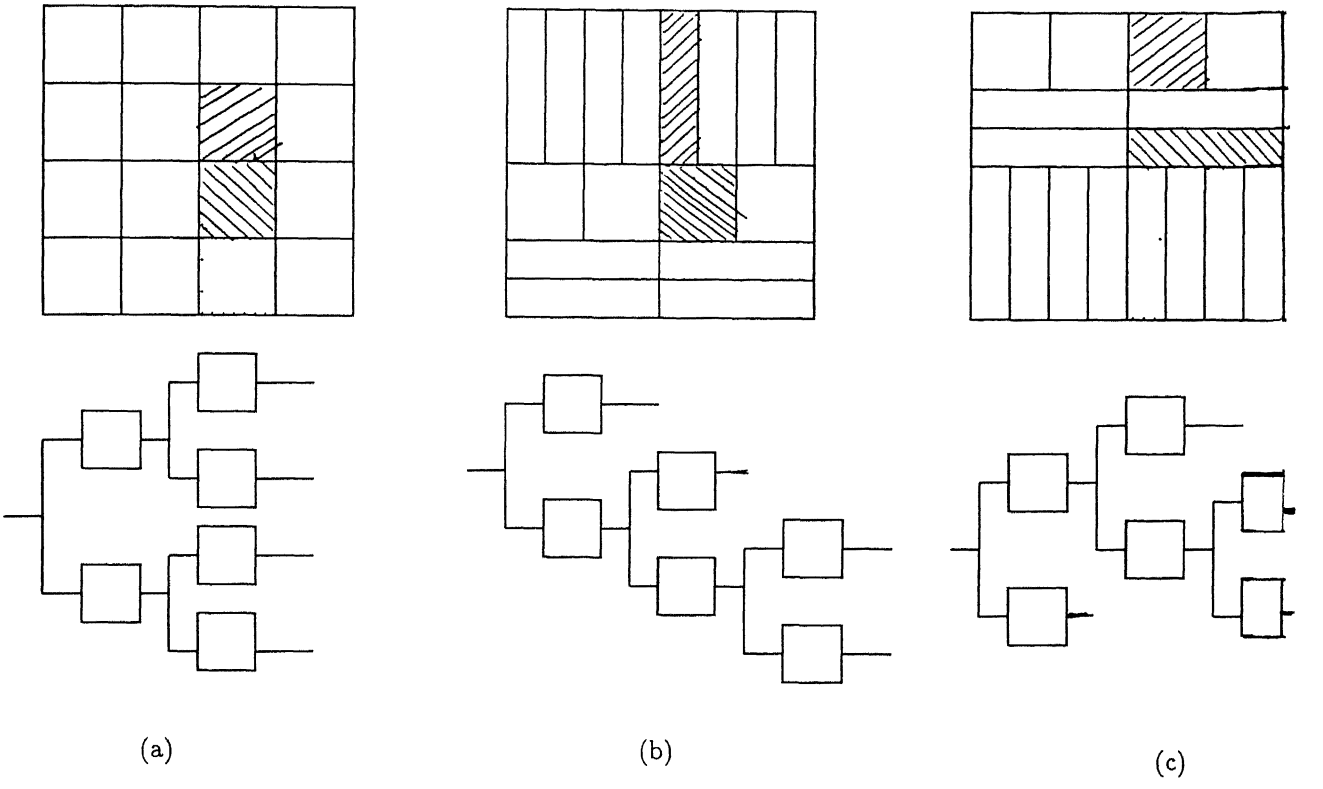


Figure 4.7: Time-frequency analysis of different filter banks (a).Full tree or STFT (b).Octave band tree or wavelet series (c).Arbitrary tree or one of the possible wavelet packet

4.1.9 Adapted subband coding

Wavelets decorrelate pictures which are close to self similar. It is not clear that any fixed choice of subbands will contain suitable templates but it is possible to use a library of bases of wavelet packets which are efficiently encoded superpositions of wavelets. These adapted subband bases come with a natural quadtree organization and some remarkable orthogonality properties.

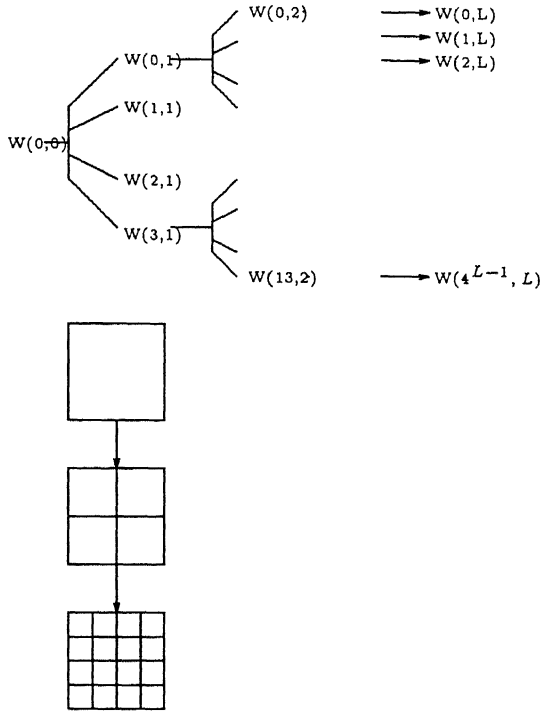


Figure 4.8: Quadtree decomposition of wavelet packets

A large library of adapted subband bases can be limited by retaining all amplitudes in the quadtree. The amplitudes produced at each stage are correlations of the signal with compactly supported oscillatory functions called wavelet packets. From the tree W of subspaces we may choose a basis subset defined as a collection of mutually orthogonal subspaces $W \in W$, or lists of pairs (n,m) which together span the root. For 2-D image various decomposition into subbands is shown in fig 4.9.

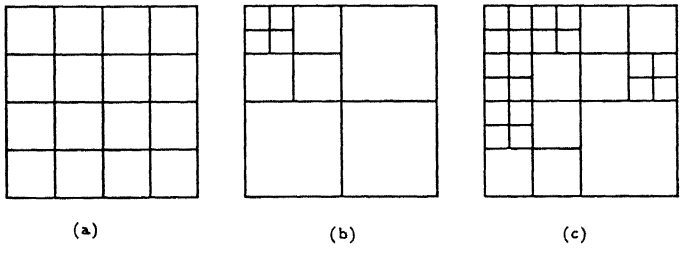


Figure 4.9: (a).Basis of subbands (linear) at one level (b).Multiresolution or wavelet basis (c).Adapted subband or wavelet packet basis

It is possible to introduce a cost function and pick a "best" wavelet basis.

A map μ from sequence x_i to R is called an additive information cost function if $\mu(0) = 0$ and $\mu(x_i) = \sum_i \mu(x_i)$

Motivated by ideas from signal processing and communication theory we were led to measure the "distance" between a basis and a function in terms of the Shannon entropy of the expansion. More generally, let H be a Hilbert space, let $v \in H, ||v|| = 1$ and assume that H is an orthogonal direct sum $H = \oplus \sum H_i$ We write $v = \oplus \sum_i v_i$ for the decomposition of v into its H_i components and define

$$\epsilon^2(v; H_i) = -\sum ||v_i||^2 ln||v_i||^2 \tag{4.36}$$

as a measure of distance between v and the orthogonal decomposition. ϵ^2 is characterized by the Shannon equation, which is a version of Pythagorus' theorem. Let

$$H = \oplus(\sum H^i) \oplus (\sum H_j) \tag{4.37}$$

$$= H_+ \oplus H_- \tag{4.38}$$

Thus H^i and H_j give orthogonal decompositions.

$H_+ = \sum H^i, H_- = \sum H_j$.Then

$$\epsilon^2(v; H^i, H_j) = \epsilon^2(v; H_+, H_-) \tag{4.39}$$

$$= ||v_+||^2 \epsilon^2(\frac{v_+}{||v_+||}; H^i) + ||v_-||^2 \epsilon^2(\frac{v_-}{||v_-||}; H_j) \tag{4.40}$$

That is the Shannon equation for entropy, if we interpret $||P_{H_+} v||^2$ to be, as in quantum mechanics, the "probability" of v being in the subspace H_+. This equation enables us to search for a smallest entropy spatial decomposition of a given vector.

The following algorithm finds the best basis based on entropy. Let

$$H_w(S) = -\sum_{x \in S_w} x^2 log x^2 \tag{4.41}$$

where $H_w(S)$ measures the expense of including W in the decomposition used to represent the picture S.

Define the best basis for representing S to be the basis subset B_0 which minimizes $\sum_{w \in B} H_w(S)$ over all basis subset $B \subset W$.

For a predetermined deepest level L, label as "kept" each subspace at level L, i.e. the subspace indexed by (n,L) for $0 \leq n < 4^L$. Next, set the level index m to L-1. Compare the information cost of the subspace W(n,m) with the sum of the information costs of its children W(4n,m+1), W(4n+1,m+1), W(4n+2,m+1), W(4n+3,m+1). If the parent is less than or equal to the sum of the children, then mark the parent as "kept". This means that by choosing the parent rather than children, we will have less entropy in the representation of S. On the other hand if the sum of the children is less than the parent, leave the parent unmarked but attribute to the sum of children's information costs. By passing this along, prior generations will always have their information costs compared to the least costly collection descendants.

After all the subspaces at level m=L-1 have been compared to their children, decrement the level index and continue the comparison. At each level we are comparing the information cost of a node to the sum of the lowest information costs obtainable by any decompositions of its four children. We can proceed in this way until we have compared the root W(0,0) to its four children. It is claimed that the topmost "kept"

nodes in depth first order constitutes a best basis.

4.2 Zerotree Coding

4.2.1 Embedded coding

An embedded code [25] represents a sequence of binary decisions that distinguish an image from the null, or all gray,image. Since, the embedded code contains all lower rate codes embedded at the beginning of the bit stream, effectively, the bits are placed in order of importance. Using an embedded code, an encoder can terminate the encoding at any point thereby allowing a desired bit rate to be met exactly. When the desired bit rate is met, the encoding simply stops. Similarly, given a bit stream, the decoder can cease decoding at any point and can produce reconstruction corresponding to all lower rate encodings.

4.2.2 Zerotree coding of wavelet coefficients

In a hierarchical subband system, with the exception of the highest frequency subbands, every coefficient at a given scale can be related to a set of coefficient at next finer scale. The coefficient at the coarse scale will be called the "parent" node, and all the coefficient corresponding to the same spatial or temporal location at the next finer scale are called "child" nodes. For a given "parent" node, the set of coefficients at all finer scales corresponding to the same location are called "descendants". Similarly, for a given child , the set of coefficient at all coarser scales of similar orientation corresponding to the same location are called "ancestors".

For a QMF-pyramid subband decomposition , the parent child dependencies are

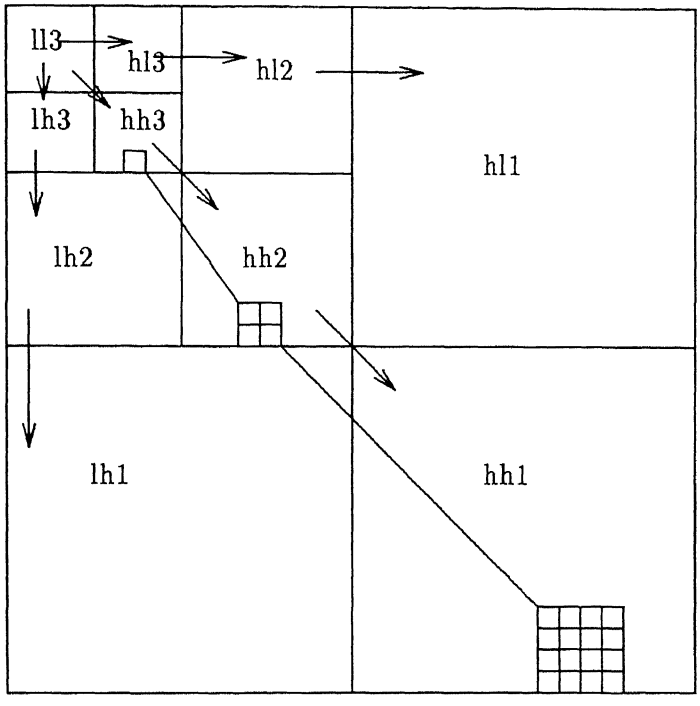


Figure 4.10: Parent -child dependencies of subbands:Note that the arrow points from the subbands of parent to the subbands of the children. The lowest frequency subband is the top left, and the highest frequency subband is bottom right. also shown is a wavelet tree consisting of all of the descendants of a single coefficient in hh3. The coefficient in hh3 is a zerotree root if it is insignificant and all of its descendants are insignificant.

shown in 4.10. With the exception of the lowest frequency subband, the parent child relationship is defined such that each parent node has three children. The scanning of the coefficient is performed in such a way that no child node is scanned before its parent. For a N-scale transform , the scan begins at lowest frequency subband, denoted as LL_N , and scans subbands HL_N ,LH_N, HH_N at which point it moves on the scale N-1,etc. The scanning pattern for a 3-scale QMF-pyramid is shown in figure 4.11. Note that each coefficient within a given subband is scanned before any coefficient in the next subband.

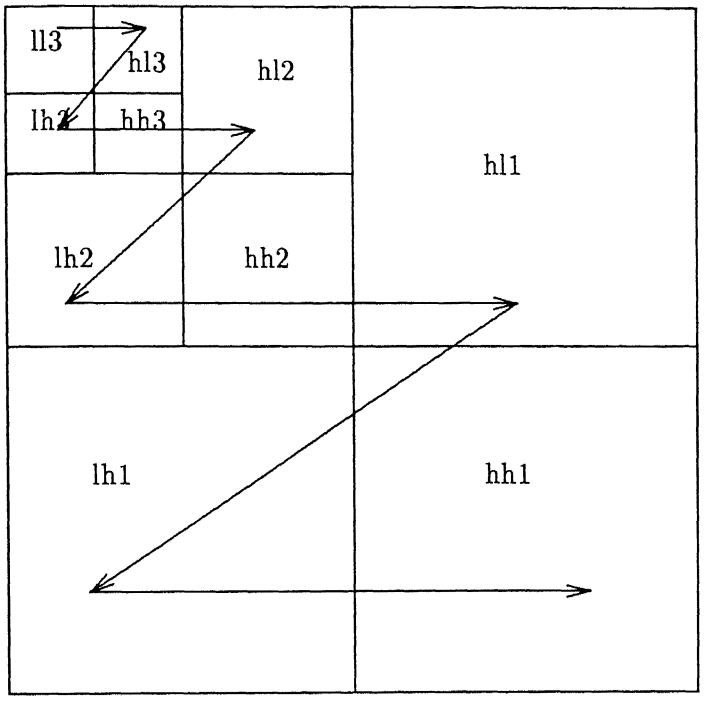


Figure 4.11: Scanning order of the subbands for encoding a significant map. Note that all positions in a given subband are scanned before the scan moves to the next subband.

4.2.3 Zerotree data structure

This data structure has four symbols.

1.Positive significant (POS) : The symbol POS is coded when the coefficient is positive and the magnitude is more than the threshold value.

2.Negative significant (NEG) : The symbol NEG is coded when the coefficient is negative and the magnitude is more than the threshold.

3.Isolated zero (IZ) : The symbol IZ is coded when the magnitude of the coefficient is insignificant, i.e. less than the threshold, but has atleast one significant descendant.

4.Zerotree root (ZTR) : The symbol ZTR is coded when the coefficient and all its descendants are insignificant.

4.2.4 Compression of significance maps using zerotrees of wavelet coefficients

The zerotree is based on the hypothesis that a wavelet coefficient at a coarser scale is insignificant with respect to a given threshold T, then all wavelet coefficients in the same spatial location at finer scales are likely to be insignificant with respect to T. Empirical evidence suggests that this hypothesis is true. When encoding the finest scale coefficients, since coefficients have no children, the symbols in the string come from a 3-symbol alphabet, whereby the zerotree symbol is not used.The flowchart for the decisions made at each coefficient are shown in figure 4.12.

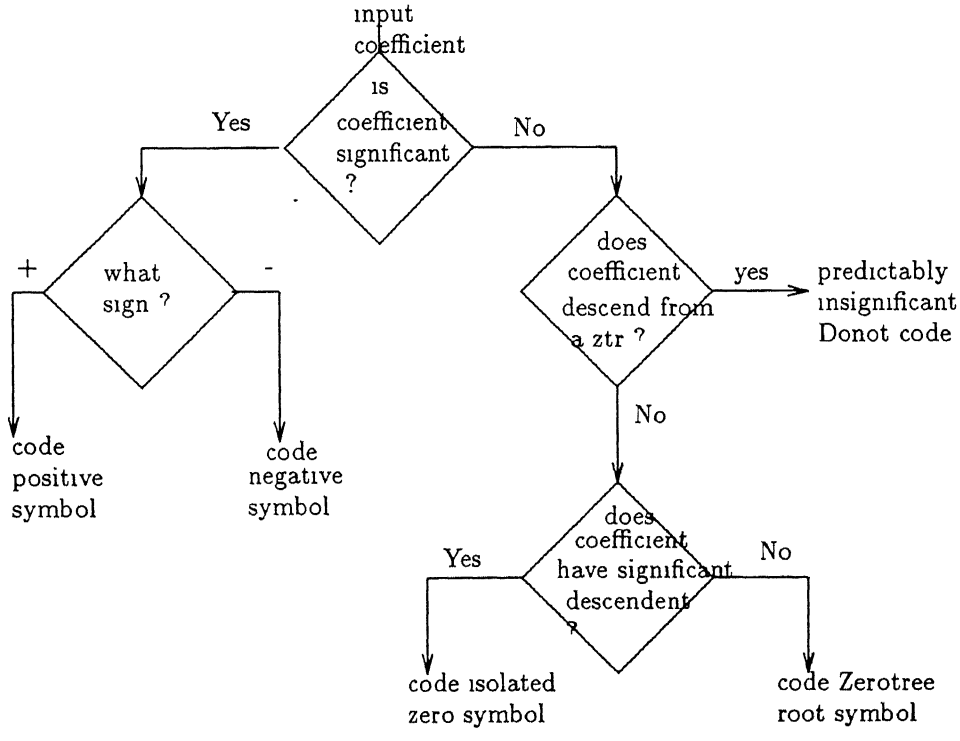


Figure 4.12: Flow chart for encoding a coefficient of the significant map

Zerotree coding reduces the cost of encoding the significant map using self similarity . Even though the image has been transformed using a decorrelating transform the

occurrences of insignificant coefficients are not independent events. More traditional techniques employing transform coding typically encode the binary map via some form of run-length coding. Unlike the zerotree symbol, which is single terminating symbol and applies to all tree depths, run-length encoding requires a symbol for each run-length which must be encoded A technique that is closer in spirit to the zerotrees is the end-of-block (EOB) used in JPEG, which is also a terminating symbol indicating that all remaining DCT coefficient in the block are quantized to zero. To see why zerotree may provide an advantage over EOB symbols, consider that a zerotree represents the insignificance information in a given orientation over an approximately square spatial area at all finer scales upto and including the scale of the zerotree root. Because the wavelet transform is a hierarchical representation, varying the scale in which a zerotree root occurs automatically adapts the spatial area over which the insignificance is represented. The end of block (EOB) symbol ,however, always represents insignificance over same spatial area, although the number of frequency bands within this spatial area varies. Given a fixed block size, such as 8x8, there is exactly one scale in the wavelet transform in which if a zerotree root is found at that scale, it corresponds to same spatial area as the block of the DCT. If a zerotree root is found at the coarser scale, then the insignificance pertaining to that orientation can be predicted over a large area.Zerotree approach can isolate interesting nonzero details by immediately eliminating large insignificant region from consideration. In this zerotree approach the focus is on reducing the cost of encoding the significant map so that, for a given bit budget, more bits are available to encode expensive coefficients. In practice , a large fraction of the insignificant coefficients are efficiently coded as part of the zerotree.

4.2.5 Zerotree like structures in DCT and wavelet packets

The concept of predicting the insignificance of coefficients from low frequency to high frequency information corresponding to the same spatial localization is a fairly general concept and not specific to the wavelet transform configuration. Zerotrees are equally applicable to quincunx wavelets, in which case each parent would have two children instead of four except for the lowest frequency, where parents have a single child.

Similar approach can be applied to linearly spaced subband decomposition such as the DCT, and to more general subband decompositions such as wavelet packets and Laplacian pyramids. For example one of the possible parent-child relationship for linearly subbands can be seen in fig 4.13.

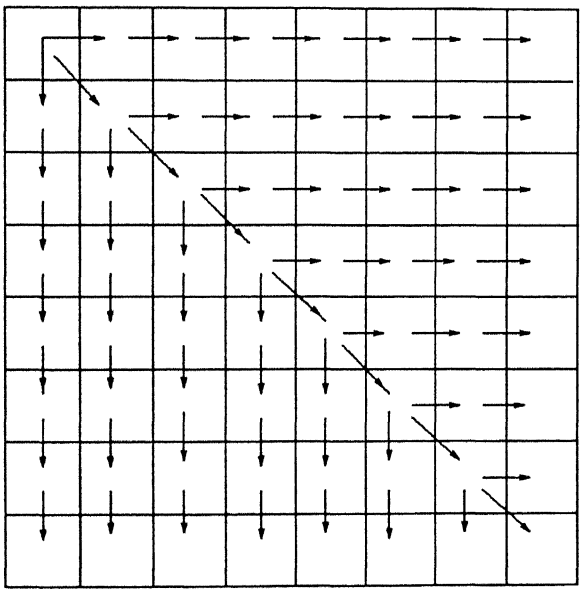


Figure 4.13: One of the possible parent-child dependencies in linear spaced subbands systems such as the DCT. Note that the arrow points from the subband of the parents to the subband of the children.The lowest frequency subband is the top left, and the highest frequency subband is bottom right.

With the use of linearly spaced subbands, zerotree like coding loses its ability to

adapt the spatial extent of the insignificance prediction. Nevertheless it is possible for zerotree like coding to outperform EOB coding since more coefficients can be predicted from the subband along the diagonal

For the case of wavelet packets, the situation is more complicated, because a wider range tiling of the "space-frequency" domain are possible. In that case, it is not always possible to define similar parent-child relationships because a high frequency coefficient may in fact correspond to a larger spatial area than a co-located lower frequency coefficient.As can be seen from figures 4.7 & 4.9 ,where arbitrary tree is used as the basis, in which it is clear that high frequency coefficient is corresponding to more spatial area than the low frequency coefficient.The best basis of the wavelet packet library is also an arbitrary tree subband decomposition, so it is not possible to define similar parent-child relationship.

It can be seen as zerotrees are predicting the insignificance of coefficients from low frequency to high frequency information.This prediction can be applied in arbitrary subband configuration as in the case of linear spaced subband discussed above, of course this does not exploit the spatial adaptivity. Considering zerotrees as merely a prediction technique, it is found possible to use zerotrees data structure for coding arbitrary tree decomposed or more specifically the bestbasis decomposed images. That is irrespective of what the subband decomposition is , high frequency coefficients are being predicted in the same way as in the wavelet transform decomposition case. This scheme is not as efficient as in the wavelet transform cofiguration but for the bestbasis case this gives almost the equivalent performance.

Summarizing it can be said as it is possible to define zerotree structure for other subband decomposition but it is most efficient for the case of wavelet transform decomposition.

4.2.6 Successive approximation quantization

In the previous section we described a method of encoding significance maps of wavelet coefficients that, at least empirically, seems to consistently produce a code with a lower bit rate than either the empirical first order entropy, or a run-length code of the significant map.

To perform the embedded coding ,successive approximation quantization is applied. The successive approximation quantization sequentially applies a sequence of thresholds $T_0, T_1, ..., T_{N-1}$ to determine significance, where the thresholds are chosen so that $T_t = T_{\frac{t-1}{2}}$.

A wavelet coefficient x is said to be insignificant with respect to a given threshold T if $|x| < T$.The initial threshold T_0 is chosen so that $|X_j| < 2T_0$ for all transform coefficient x_j.

During the encoding (and decoding) two link-lists, one for dominant pass and the other for subordinate pass, of wavelet coefficients are maintained. At that point in the process ,the dominant list keeps track of the coefficients that have not found to be significant in the same relative order as the initial scan. This scan is such that the subbands are ordered, and within each subband, the set of coefficients are or-

dered. Thus using the ordering of the subbands shown in figure 4.11, the subordinate list contains the magnitudes of those coefficients that have been found to be significant.

During the dominant pass coefficients that have not been found to be significant in the previous scan, are compared to the threshold T_i to determine the significance, and if significant ,their sign. The significance map is then zerotree coded. Each time a coefficient is encoded as significant, (positive or negative), its magnitude is appended to the subordinate list.The coefficient that had been determined to be significant in the previous scan is considered as insignificant in all the following dominant passes at smaller thresholds.

During a subordinate pass, the width of the effective quantizer step size, which defines an uncertainty interval for the true magnitude of the coefficient, is halved. For each magnitude on the subordinate list, this refinement can be encoded using binary alphabet with a 1 symbol indicating the lower half. The string of symbols from this binary alphabet that is generated during a subordinate pass is then entropy coded.

This process continues alternately between dominant passes and subordinate passes where the threshold is halved before each dominant pass.

In the decoding operation, each decoded symbol, both during a dominant and a subordinate pass, refines and reduces the width of the uncertainty interval in which the true value of the coefficient (or coefficients in the case of a zerotree root) may occur. the reconstruction value can be any where in that uncertainty interval. For minimum mean square error distortion, one could use the centroid of the uncertainty region using

some model for the PDF of the coefficients. However a practical approach, that is used in the experiments , which is also MINMAX optimal, is to simply use the center of the uncertainty interval as the reconstruction value.

The encoding stops when some desired terminating condition is met, such as when the bit budget is exhausted.

4.3 Arithmetic Coding

Arithmetic coding is a lossless compression technique that produces an encoded string for an input string of symbols and model. This encoded string represents a fractional value R for the range $0 <= R < 1$.

Arithmetic coding [29] is superior to the well known Huffman method in many respects. It represents information as compactly as Huffman code. It is known that each symbol in the input string is represented as an integral number of bits in the encoding, then Huffman coding achieves "minimum redundacy". In other words , it performs optimally if all symbol probabilities are integral powers of 1/2. But in practice this is not so. Arithmetic coding dispenses with this restriction that each symbol be represented as an integral number of bits.

In arithmetic coding, a message is represented by an interval of real numbers between 0 to 1. As the message becomes longer , the interval needed to represent it becomes smaller, and the number of bits needed to specify that interval grows. Succes-

sive symbols of the message reduce the size of interval in accordance with the symbol probabilities generated by the model. The more likely symbols reduce the range by smaller amounts as compared to unlikely symbols and add fewer bits to the message. Before transmitting a message the range of the message in the entire interval is [0,1), and the half open interval is denoted by $0 <= x < 1$.As each symbol is processed, the range is narrowed to the portion allocated to the symbol.

We can divide models for the arithmetic coding into two categories. 1.Fixed model and 2.Adaptive model

In the fixed model, frequency of symbol occurences are taken from sample text.

In the adaptive model , frequencies are initialized to some value during encoding, and these frequencies are updated on the basis of symbol frequencies observed in the input string.

Arithmetic encoder operates successively on each data symbol, determines the context (i.e. which relative frequency distribution applies to the current event) and generates the code string.

To represent the magnitude R of the encoded string in the interval $[0, 1)$ great precision is required. Fortunately this magnitude need not be given all at once. At any stage the upper and lower bounds for R are available as a finite no. of digits. These digits are left shifted as they become identical and new digits are brought at the low-significant end.

4.3.1 Arithmetic coding in context of zerotrees

Note that the particular alphabet used by the arithmetic coder at any given time contains either 2,3 or 4 symbols depending on whether the encoding is for subordinate pass , or dominant pass with no zerotree root symbol, or a dominant pass with a zerotree root symbol. There is advantage in adapting the arithmetic coder. Since there are never more than four symbols, all the possibilities typically occur with in a reasonably measurable frequency. This allows an adaptation algorithm with a short memory to learn quickly and track continuously changes in symbol probabilities. This adaptivity accounts for some of the effectiveness of the overall algorithm. On the other hand in case of algorithms that donot use succesive approximation, several events are needed before an adaptive entropy coder can reliably estimate the probabilities of unlikely symbols.

Once the type of model (adaptive model for the present case) is fixed for arithmetic coding,maximum frequency count is the critical parameter for the coding, because it affects underflow, overflow and learning rate for adaptation. Arithmetic coding works by scaling the cumulative probabilities given by the model onto the interval [low, high] for each character encoded.If they are very close together , then there is possibility of mapping different symbols in the same interval. Therefore , the interval should atleast be as large as possible. It should not be too large to cause overflow. Learning rate for adaptation is inversely proportional to the maximum histogram count. Again, very small symbol set from zerotree coding is adavantageous in choosing maximum histogram count. A maximum frequency count of 256 is used in taking account all the

factors given above.

4.4 Overhead information of spatial decorrelation subsystem

This contains only 8 bytes header.After this header entire bit stream is arithmetically encoded by a single arithmetic coder with an adaptive model. This header contains (1). No. of wavelet scales (2). Dimension of the image (3). The initial threshold (4). Mantissa.

4.5 Bit rate assignment for eigen images

Compression performance is strongly influenced by the selected bit assignment scheme for the spectrally decorrelated eigen images. Since the variances of the ordered eigen planes decrease almost exponentially , it may be suggested to code the eigen images at rates proportional to the logarithm or squareroots of their variances. Although this bit assignment strategy may result in lowest MSE, it is not suitable for multispectral imagery, the subtle variations in the spectral signatures of certain terrains manifest themselves in the lower eigen images with the same level of accuracy as the others , irrespective of their variances of the resulting overall mean square coding error . In general the criterion for bit assignment should induce a uniform coding error on all the reconstructed eigen planes.

These results are derived from an extensive set of experimentation and cannot be derived from theoretical principles.

To implement the optimum bit assignment scheme, an arbitrary MSE for the first eigen image is selected. The effective MSE for the remaining eigen images are calculated as

$$MSE(n) = L(1) * \frac{B(n)}{B(1)}, n = 2, 3, 4..... \tag{4.42}$$

Since it is not known in advance what MSE results for a particular eigen image at given bit budget, it is necessary to look-up table of MSE vs. bytes be transmitted for each eigen image.

In general it is beneficial to replace some of the low variance eigen images by their mean values. This results in (1). good performance in terms of MSE (2). considerable savings in computation and power requirement, since fewer eigen images need to be coded.

The most simple method is a compromise between complexity and performance. In which bit rate is assigned to eigen images based on variance and the initial quantization error.

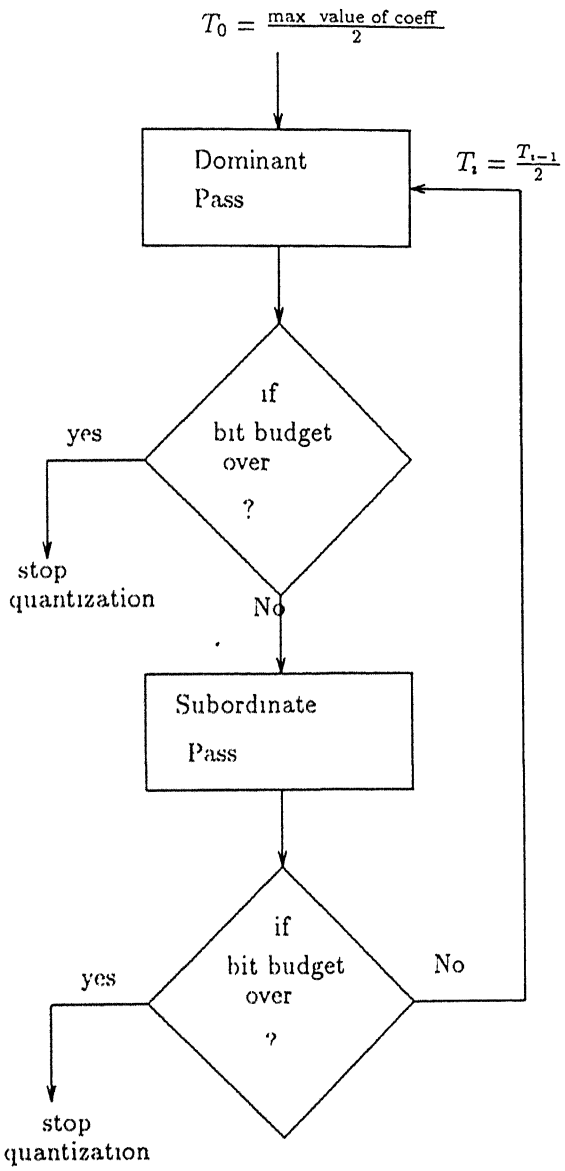


Figure 4.14: Successive approximation quantization flowchart

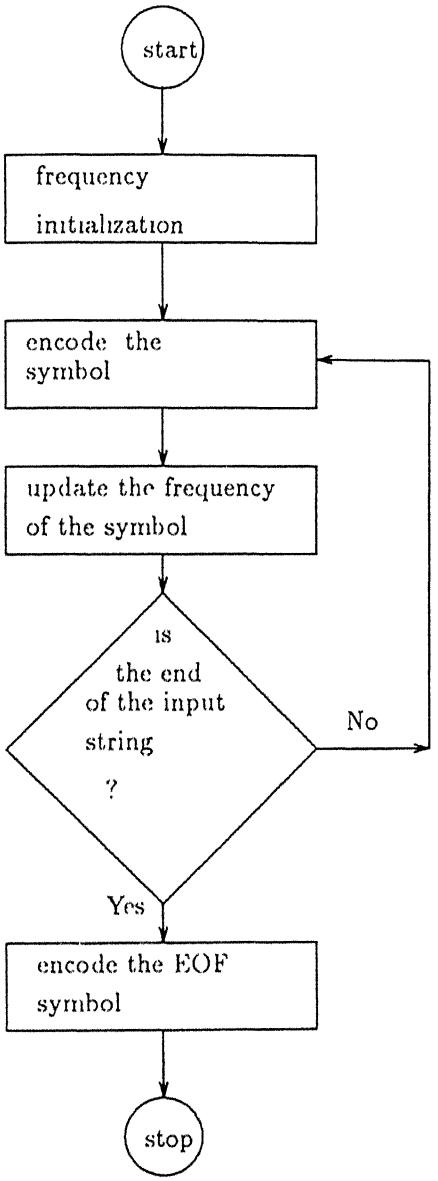


Figure 4.15: Encoder for arithmetic coding

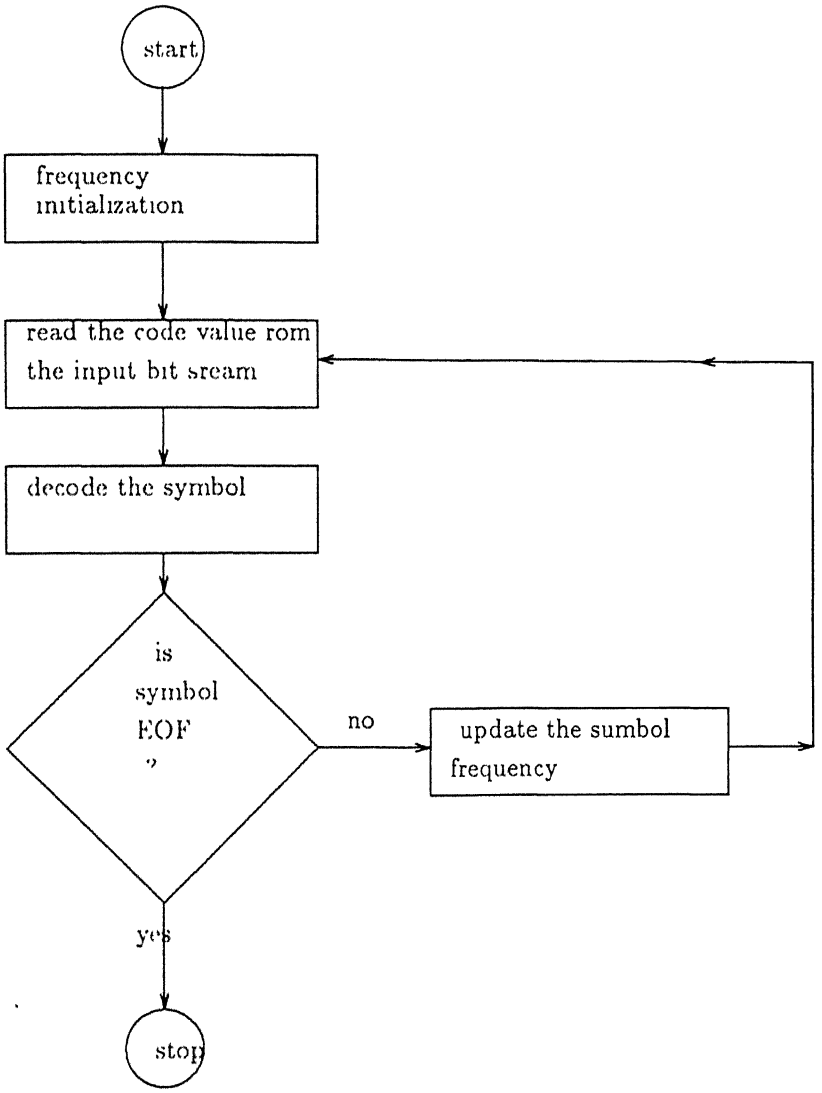


Figure 4.16: Decoder for arithmetic coding

Chapter 5

Experimental Results and Discussion

The present chapter discusses the experimental results of the following multispectral image compression schemes :

(1).KL transform followed by embedded zerotree coding using wavelet transform

(2).KL transform followed by embedded zerotree coding using wavelet packet transform coefficients.

Each of these schemes were examined using two different sets of images. The image sets used are :

(1). A semi-urban image

(2). An urban image

Figures (img1) & (img2) show the original image sets 1 & 2 respectively. The test images contain diverse range of natural and urban terrains and as such are difficult for compression. The specifications of the test image sets 1 & 2 are given below .

(a). Spectral resolution -

Each set is a 4 band image in the following frequency range :

Band 1 0.45 - 0.52 μm

Band 2 - 0.52 - 0.6 μm

Band 3 0.6 - 0.7 μm

Band 4 - 0.7 - 0.8 μm

(b). Spatial resolution-

The spatial resolution measured in terms of ground sample distance (GSD) is 36.25m.

(c). Radiometric precision-

Each image pixel is represented by 7 bpp or 0-128 gray levels.

5.1 Measures for compression system efficiency

The correlation coefficient is a convenient and useful method to measure the inherent spectral decorrelation. The correlation coefficient matrix is defined as normalized covariance matrix. That is, the correlation coefficient for each pair of bands is equal to their covariance value divided by the squareroot of the product of their variances.

Objective measures to judge the quality of the reconstructed images are peak signal to noise ratio and average mean square error. As is common in the image coding literature [11] , the square of the maximum intensity is used as the definition of signal power and the values are reported in dB units. The mean square error of a reconstructed image of size NxN is defined as

$$MSE = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} ||x_{ij} - \hat{x}_{ij}||^2 \tag{5.1}$$

where x_{ij} is the pixel at the i^{th} row and j^{th} column in the image and $\hat{x}_{ij}$ is its reconstructed value.

The power ratio for 7 bpp image set is given by

$$\rho = \frac{127^2}{MSE} \tag{5.2}$$

The value in dB units of PSNR is given by

$$PSNR = 10log_{10}\rho \tag{5.3}$$

The bit rate for the image set is given by

$$bitrate(bpp) = \frac{\text{Total no. of bytes transmitted} * 8}{\text{Total no. of pixels}} \tag{5.4}$$

5.2 Spectral Decorrelation Efficiency

It is well known that KL transform is an optimum method to spectrally decorrelate the multispectral images. Figures (img3) and (img4) show the spectrally decorrelated eigen images of test image set 1 & 2 respectively. The compaction of the data as a result of KLT operation is clearly evident. It is observed that more than 90 % of the energy or information content of the data sets resides in the first two eigen images and more than 98 % of the energy in the three eigen images. The remaining eigen image has very little information content of the data and as such require substantially fewer bits to be coded.

A convenient approach to measure the amount of data compaction is to measure variances of the eigen images. The variance of an image reflects its busyness or information content.

Figures 5.1, 5.2 and 5.3, 5.4 show the covariance and correlation coefficient matrix of image set 1 and image set 2 respectively and figures 5.5, 5.6 and 5.7, 5.8 show the

covariance and correlation coefficient matrix for spectrally decorrelated eigen images.

10 14	6 99	16 96	6.98
6.99	5,77	13 6	7 79
16.96	13.6	34.5	15.5
6.98	7.79	15.5	82 79

Figure 5.1: Covariance matrix of original semi-urban image set

1	0.91	0.9	0.24
0.91	1	0.96	0.35
0.9	0.96	1	0.29
0.24	0.35	0.29	1

Figure 5.2: Correlation coefficient matrix of original semi-urban image set

28.43	22 26	43.56	4.29
22.26	21.19	41.78	8.4
43.56	41.78	84.67	16.12
4.29	8.4	16.12	46.08

Figure 5.3: Covariance matrix of original urban test image

1	0.9	0.88	0.11
0.9	1	0.98	0.27
0.88	0.98	1	0.26
0.11	0.27	0.26	1

Figure 5.4: Correlation coefficient matrix of original urban image set

90 93	0 11	-0.19	0 19
0.11	40.54	-0.4	0.15
-0.19	-0 4	1.034	0.109
0 19	0.15	0.109	0 79

Figure 5.5: Covariance matrix of eigen images of semi-urban test image

It is observed that the off-diagonal values of correlation coefficient matrix of eigen images are zero. That is, the eigen images are completely decorrelated. The variance of eigen images drops almost exponentially with the order of the eigen images. Plots 1 and 2 show the normalized variance with respect to bands no. for image sets 1 and 2 respectively. The steeper the drop in the variance in moving from low to higher order eigen images, more efficient data compaction and spectral decorrelation is achieved.

To compare the performance of KLT with other spectral decorrelation schemes , experiments are done on these test images using DCT and affine prediction . Tables 5.1 & 5.2 show the variation of normalized variance with respect to band no. for KLT,

band no.	variance
1	90.43
2	40.54
3	1.034
4	0.79

(a)

band no.	variance
1	50.92
2	39.35
3	31.82
4	11.38

(b)

band no.	variance
1	84.07
2	13.48
3	10.14
4	1.47

(c)

Table 5.1: Comparison of spectral decorrelation efficincy of different spectral decorrelation techniques.The tables show the variance of eigen planes (a).when KLT is used (b).when DCT is used (c).when affine predictor is used to spectrally decorrelate the original image set 1

band no.	variance
1	132.678
2	43.26
3	3.58
4	0.60

(a)

band no.	variance
1	112.944
2	27.32
3	22.32
4	17.87

(b)

band no.	variance
1	98.564
2	31.36
3	28.13
4	4.26

(c)

Table 5.2: Comparison of spectral decorrelation efficiency of different spectral decorrelation techniques.The tables shows the variance of eigen planes 1).when KLT is used 2).when DCT is used 3).when affine predictor is used to spectrally decorrelate the original image set 2

1	0.0	0.0	0.02
0.0	1	0.06	0.02
0.0	0.06	1	0.10
0.02	0.02	0.10	1

Figure 5.6: Correlation coefficient matrix of eigen images of semi-urban test image

105.38	0 002	-0 053	0.05
0 002	23.79	-0.015	-0 023
-0.053	-0.015	3.023	-0.002
0.05	-0.023	-0.002	0 49

Figure 5.7: Covarince matrix of eigen images of urban test image

DCT and affine prediction spectral decorrelation schemes for test images 1 & 2 respectively. It is observed that for the KLT scheme two eigen images are having negligible variance whereas in DCT and affine predictor scheme only one band is having negligible variance. Referring to the plots 1 and 2, which also show the variation of normalized variance with respect to band no. for the DCT and affine predictors, it is observed that the KL transform is the most efficient of all other transforms.

Although KL transform has no fast algorithm for its computation and requires few overhead bits , it is the optimum decorrelating transform. Overhead bits needed is of the order of 0.0003 bits per pixel which is negligible in comparison with the transmit-

1	0.0	0.0	0.0
0.0	1	0.0	0.0
0.0	0.0	1	0.0
0.0	0.0	0.0	1

Figure 5.8: Correlation coefficient matrix of eigen images of urban test image

ting bits (0.09 bpp) per pixel. Affine prediction may work better if prediction is done, not only with feature band alone but with all other bands.

In the present work, test image sets used have only 4 bands. From the results reported in [24] , it is concluded that the spectral decorrelation efficiency increases with the no. of bands. Hence better compaction can be expected on using multispectral image data sets containing more no. of bands.

5.3 Terrain adaptive test

The terrain adaptive test was performed on these test images.Plot 3 shows the relative sizes of variance of each eigen plane of image set 1 for non-terrain adaptive and terrain adaptive approaches for different window size of 128, 64 and 32 respectively. The significant decrease in the variance as a result of terrain adaptation is observed, which indicates a greater compaction of data.

From the fig (img5) which shows the eigen images for the terrain adaptive approach

with a window size 64, it is observed that for the terrain adaptive approach, the eigen images appear to have discontinuities over the edges of the selected covariance update window. The blocking effect clearly indicates the adaptation of the spectral transformation process to the characteristics of the terrain within the selected window.

Table 5.3 gives the rate vs. distortion results when terrain adaptive approach with a window size of 64 is used. In plot 4 a comparison of mean square error vs. bpp of non-terrain adaptive and terrain adaptive approaches is made when each of them is followed by embedded zerotree coding of wavelet coefficients. It is observed that terrain adaptive approach does not improve the performance (mse vs. bpp). Fig. (img6) shows the reconstructed images of test image set 2 at 80:1 CR (i.e. 0.09 bpp). It is observed that blocking effect is visible even in the reconstructed image.

5.4 Overhead in spectral decorrelation subsystem

In the present work only linear quantization was used to map eigen planes to eigen images as the dynamic range was low (less than 127) . For the non-terrain adaptive approach overhead bits B_{oh} needed is obtained using results given in section 3.4

N = No. of bands =4

B = Image dimension = 512

$$B_{oh} = \frac{40N + 8N^2}{NB^2} bpp$$
$$= \frac{40 * 4 + 8 * 4^2}{4 * 512 * 512} bpp$$
$$= 0.000274 bpp$$

bpp	CR	mse	PSNR
0.7	10 : 1	3.52	36.61
0.44	16 : 1	4.37	35.67
0.32	22 : 1	5.15	34.95
0.18	40 : 1	7.43	33.36
0.11	62 : 1	9.37	32.35
0.09	77 : 1	10.22	31.98

Table 5.3: Rate vs. distortion results when Terrain adaptive KLT with window size 64 is followed by zerotrees in wavelet coefficients

It can be seen that B_{oh} is quite insignificant in comparison to bits per pixel tansmitted.

For terrain adaptive approach it is given by

$$B_{oh} = K^2 * \text{overhead of non} - \text{terrain adaptive approach} \tag{5.5}$$

when

$$K = \frac{\text{dimension of image}}{\text{dimension of window}} \tag{5.6}$$

for window size 64 and image dimension 512 overhead comes to 0.017 bpp. It is quite significant at very low bit rate i.e. 0.09 bpp.

5.5 Non-terrain adaptive test

The non-terrain adaptive compression algorithms have been simulated for two sets of test images. The performance measure used for the compression algorithms is rate vs. distortion variation. The performance of the different compression algorithms are given in tables 5.4,5.5 and 5.6 respectively for test image set 1 and in tables 5.7,5.8 and 5.9 respectively for test image set 2.

Table 5.4 gives the estimate of performance when KL transform is used with embedded zerotree root coding using wavelet transform coefficients. Table 5.5 shows the performance when KLT is followed by embedded coding using arbitrary wavelet packet coefficients (i.e. linear subband decomposition when all the coefficients of lowest subbands are "kept") . Table 5.6 shows the performance when KLT is followed by zerotree of bestbasis coefficients of the wavelet packet library. Plots 5 and 6 compare the mean square error and PSNR vs. bpp of these algorithms respectively . As the variation of

mse and PSNR vs. bpp respectively for each of the above schemes is very small , it can be concluded that the performance of algorithms are comparable.

Visual evaluation of the reconstructed image set 1 (semi-urban) shown in fig (img 7) at compression ratio 100:1 (i.e. 0.07 bpp) reveals that a visually lossless performance can be achieved upto this CR.

For image set 2 (i.e. urban image) table 5.7 summarises the performance of compression when KL transform is followed by embedded zerotree coding using wavelet transform coefficients. Table 5.8 summarises the performance when KLT is followed by embedded coding using arbitrary wavelet packet coefficients (i.e. linear subband decomposition when all the coefficients of lowest subbands are "kept") and table 5.9 shows the performance when KLT is followed by embedded zerotree coding of bestbasis coefficients of the wavelet packet library. Plots 7 and 8 compare the mean square error and PSNR vs. bpp of these algorithms respectively. As the variation of mse and PSNR vs. bpp respectively for each of the above schemes is very small , it can be concluded that the performance of algorithms are comparable.

Figures (img8) , (img9) and (img10) shows the reconstructed image for set 2 at 80:1 CR , when each of the above mentioned compression algorithms are applied respectively. Visual evaluation shows that the images are visually lossless upto a compression ratio of 80:1 (i.e. 0.09 bpp).

The comparison of these algorithms show that no significant improvement is obtained by using the best basis coefficients of wavelet packet library. This is so because

bpp	CR	mse	PSNR
0.5	14 : 1	1.914	39.25
0.44	16 : 1	1.99	39.08
0.36	20 : 1	2.195	38.66
0.3	23 : 1	2.38	38.31
0.198	35 : 1	2.96	37.36
0.152	46 : 1	3.415	36.74
0.107	66 : 1	3.98	36.07
0.076	92 : 1	4.66	35.39
0.061	115 : 1	5.15	34.95
0.054	130 : 1	5.432	34.72
0.023	300 : 1	7.27	33.53

Table 5.4: Rate Vs. distortion results for semi-urban image when KLT/Zerotrees in wavelet coefficients are used

bpp	CR	mse	PSNR
0.549	12 : 1	1.72	39.72
0.5	14 : 1	1.914	39.25
0.36	20 : 1	2.175	38.77
0.27	25.5 : 1	2.36	38.33
0.168	12 : 1	2.98	37.33
0.091	76.5 : 1	4.28	35.16
0.0.045	152.5 : 1	5.39	34.75
0.0.016	448 : 1	8.19	32.93

Table 5.5: Rate Vs. distortion results for semi-urban image when KLT/Zerotrees in coefficients of best basis of wavelet packet library are used

bpp	CR	mse	PSNR
0.549	12 : 1	1.67	39.83
0.5	14 : 1	1.93	39.22
0.36	20 : 1	2.135	38.78
0.27	25.5 : 1	2.45	38.17
0.168	42 : 1	3.23	36.97
0.106	66 : 1	3.96	36.09
0.076	92 : 1	4.65	35.4
0.045	152.5 : 1	5.713	31.5
0.023	300 : 1	7.24	33.47

Table 5.6: Rate Vs. distortion results for semi-urban image when KLT/Zerotrees in wavelet packet coefficients are used

bpp	CR	mse	PSNR
1.0	7 : 1	2.0	39.06
0.9	8 : 1	2.2	38.1
0.793	9 : 1	2.53	38.0
0.70	10 : 1	2.83	37.56
0.595	11.8 : 1	3.4	36.76
0.435	16 : 1	4.154	35.89
0.36	19 : 1	4.79	35.27
0.24	29 : 1	6.159	34.18
0.17	42 : 1	7.614	33.25
0.09	77 : 1	10.69	31.78
0.038	184 : 1	57.8	24.45

Table 5.7: Rate Vs. distortion results for urban image when KLT/Zerotrees in wavelet coefficients are used

bpp	CR	mse	PSNR
0 91	7.7 : 1	2.013	39.03
0.762	9 : 1	2.38	38.29
0.640	10.9 : 1	2.89	37.46
0.565	12.4 : 1	3.203	37.02
0.435	16 : 1	3 94	36.11
0.36	19 : 1	4.535	35.51
0.33	21 : 1	4.77	35.28
0.23	30 : 1	6.17	34.17
0.1525	45 : 1	7.88	33.1
0.09	77 : 1	10.12	32.02
0.07	100 : 1	11.1	31.62
0.038	130 : 1	14.15	30.56

Table 5.8: Rate Vs. distortion results for urban image when KLT/Zerotrees in best

bpp	CR	mse	PSNR
0.91	7.7 : 1	2.06	38.93
0.762	9 : 1	2.51	38.0
0.640	10.9 : 1	3.03	37.25
0.565	12.4 : 1	3.3	36.88
0.435	16 : 1	1.084	35.96
0.36	19 : 1	4.7	35.35
0.33	21 : 1	1.97	35.11
0.23	30 : 1	6.1	34.00
0.1525	45 : 1	8.178	32.94
0.09	77 : 1	10.78	31.74
0.07	100 : 1	11.4	31.5
0.038	183 : 1	45.6	25.48

Table 5.9: Rate Vs. distortion results for urban image when KLT/Zerotrees in wavelet packet coefficients are used

bpp	CR	mse	PSNR
0.7	10 : 1	2.52	38.04
0.44	16 : 1	4.88	35.18
0.32	22 : 1	5.92	34.34
0.18	40 : 1	7.07	33.18
0.11	62 : 1	9.66	32.22
0.09	77 : 1	10.29	31.95

Table 5.10: Rate vs. distortion results when KLT is followed by zerotrees in DCT coefficients

the zerotree coding approach is not adapting to the spatial extent of the coeffecients [25] . It is only predicting the insignificant coefficients. The results also conclude that zerotrees for any arbitrary subband decomposition can be defined, it is however the most appropriate coding scheme only when wavelet transform is used.

To judge the improvement in performance of wavelet transform over DCT, embedded zerotree coding method was also applied on DCT coefficients of image set 2. Table 5.10 shows the rate vs. distortion measurement and plot 8 gives mean square error vs. bpp for image set 2. It is seen that the variation of mse and PSNR in this case is almost equivalent to that when wavelet transform was used.However DCT has an

additional drawback of producing blocking effects. Visual evaluation of the images in fig (img11) shows that the blocking effect dominates after 40:1 (0.18 bpp) compression ratio and the image is visually lossy. For the purpose of comparison, fig(img12) shows the reconstructed image set 2 at CR of 80:1, when zerotree coding uses DCT coefficients; blocking effects are clearly visible in the reconstructed images.

5.6 Spectral Fidelity

In the bandwidth compression of multispectral imagery , preservation of the spectral resolution across bands is as essential as preserving the spatial resolution within each band. This is necessary in order to preserve the spectral fidelity of the data , i.e. spectral signatures of the terrains. It is important that the spectral signatures should not be lost during the compression process .

To measure the spectral fidelity of multispectral imagery a useful tool is devised which is based on correlation coefficient matrix. The spectral fidelity is depicted by the measurement of the correlation coefficient matrix. Any deviation from the original correlation coefficient matrix indicates the loss in spectral fidelity. the loss of correlation coefficient matrix in urban image are shown in fig 5.9 .

The loss of correlation coefficient matrix in semi-urban (image set 2) is shown in the fig 5.10.

These experimental results indicate that the loss of spectral fidelity , as measured by the deviation from the original correlation coefficient matrix is insignificant.Even at the compression ratio of 80:1 for urban image the loss is no more than 0.07 i.e. only 7

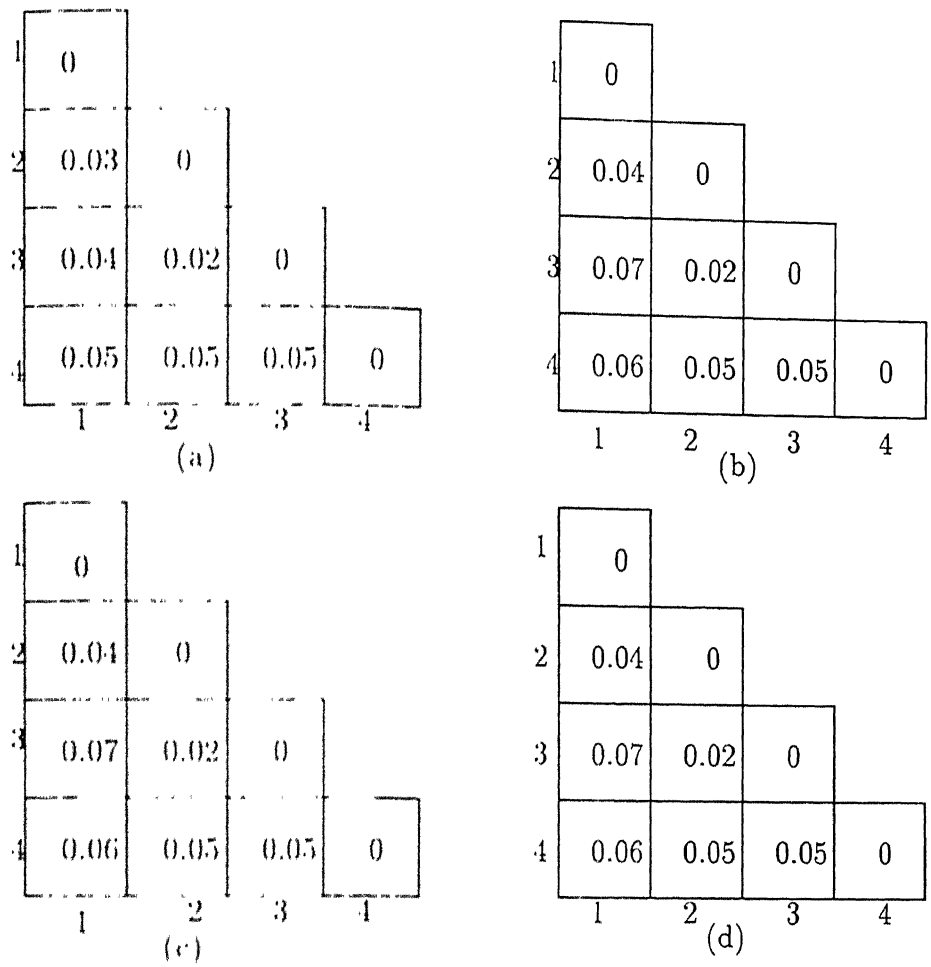


Figure 5.9: Loss in correlation coefficient matrix of image set 2 (a). when KLT is followed by zerotrees in Wavelet coefficients at CR 16:1 (b). when KLT is followed by wavelet transform at CR 80:1 (c). when KLT is followed by wavelet packet coefficients at 80:1 CR (d). when KLT is followed by best basis coefficients of wavelet packet library at 80:1 CR

percent . This loss for semi-urban image is no more than 0.13 at 100:1 CR.

For terrain adaptive case, the spectral fidelity is also calculated. For image set 2 (urban image), the loss in spectral fidelity for 80:1 compression ratio when embedded zerotree coding of wavelet coefficients is used, is given in fig 5.11.

The above calculation shows that loss in spectral fidelity is insignificant even at very low bit rate. This shows that the efficiency of algorithm is good even at very low bit rate or very high compression ratio. In terrain adaptive approach there is an

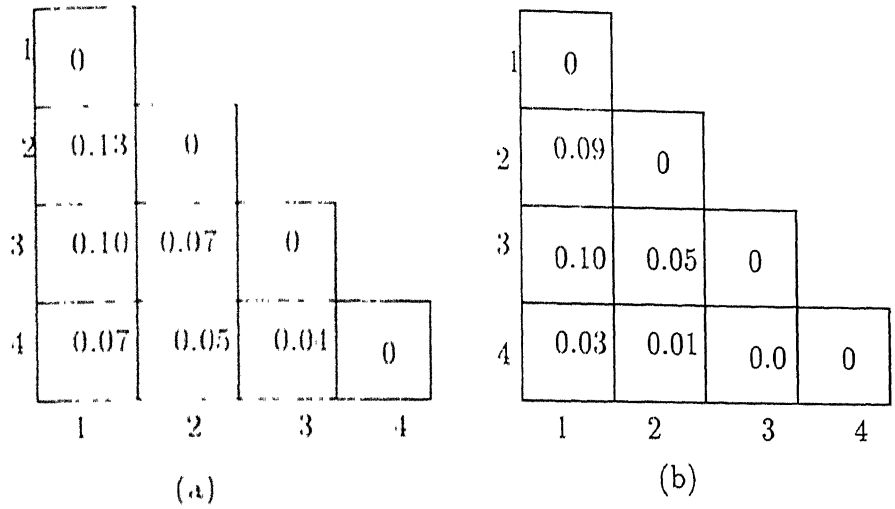


Figure 5.10: Loss in correlation coefficients of image set 1 (a). when KLT is followed by Embedded zerotree coding using wavelet coeficients at 16:1 CR (b). when KLT/zerotrees in wavelet coefficients is used at 100:1 CR

improvement of 2% in spectral fidelity between few bands. Comparing spectral fidelity for non terrain adaptive scheme (fig 5.9) and terrain adaptive scheme (fig 5.11), we observe that the improvement in spectral fidelity exists between band 4 and band 1, band 1 and band 2 & band 4 and band 3.

5.7 Comparison of performance with other compression schemes

In a recent work [28] [1] [11] related to lossy compression of multispectral imagery it is mentioned that the mean residual and gain shaped VQ yield the best results with CR 20:1. Nonlinear VQ was found to yield the best qualitative performance with a CR of 32:1.

For a hierarchical data compression system CR of 27:1 was obtained. Another system is presented [10] in which the 3-D data is first spectrally decorrelated by using

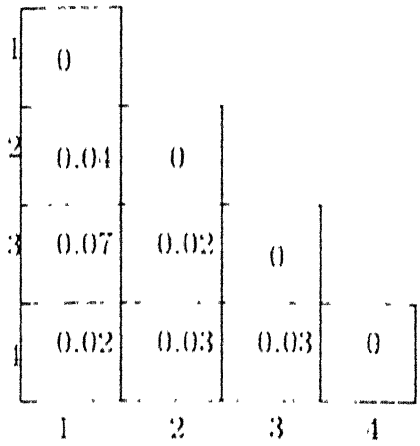


Figure 5.11: Loss in correlation coefficient matrix of image set 2 at 80:1 CR when Terrain adaptive KLT window size 64 is followed by Embedded zerotree coding using wavelet coefficients

KL transform and then spatially decorrelated the principal components using discrete wavelet transform .It was reported that this yielded 40% improvement in compression.

Recently, in [21] a compression system is described in which the KL transform is followed by JPEG. A visually lossless reconstruction of a multispectral image set with 10:1 CR was achieved.

In the present work, where KL transform is followed by embedded zerotree coding using wavelet coefficients and wavelet packet coefficients respectively gave visually lossless reconstruction for urban image at a CR of 80:1 (i.e. 0.09 bpp) .

In this scheme, encoder is fully adaptive as the model is initialized at each threshold for each of the dominant and subordinate passes. There is no training of any kind and no ensemble statistics of images are used in any way. An interesting property of embedded coding is that when the encoding and decoding is terminated during middle of a pass or in the middle of scanning of a subband there are no artifacts produced

that would indicate where the termination occur. Another interesting property of the embedded coding is that because of the implicit global bit allocation even at extremely high CR of 256:1 the image quality is poor but still recognizable.

Chapter 6

Conclusion

In this thesis an investigation of a 3 D transform based technique to compress multispectral imagery has been carried out. The algorithm exploits the inherent spectral and spatial correlation present in the image data sets. The compression technique essentially consists of two subsystems : The spectral decorrelation and the spatial decorrelation subsystem respectively. Karhunen-Loeve transform(KLT) used for spectral decorrelation of the bands was followed by an embedded zerotree coding of wavelet transform coefficients to achieve spatial decorrelation of each band.

The spectral and spatial modularity of the algorithm architecture allows EZW or KLT to be replaced by any alternate spatial or spectral coding procedure.

Superior performance of the adopted compression scheme judged from the reconstructed image fidelities ranges from near lossless at about 10:1 CR to visually lossy beginning at about 80:1 CR for two typical 4 band image data sets containing urban/semiurban terrains. Better results can be expected if test image sets with large no. of spectral bands were available. The compression achieved using the present method

is better than other multispectral image compression schemes reported in literature.

References

[1] Abousleman G.P. ,Marcellin M.W.and Hunt B R. Compression of hyperspectral imagery using the 3-D DCT and hybrid DPCM/DCT using TCQ *IEEE Trans. on geoscience and remote sensing*,Jan 1995.

[2] Antonini M. ,Barland M. P.Mathew and I.Daubechies Image Coding using Wavelet transform *IEEE Trans. on Image Processing*,Apr 1992.

[3] Burt P.J. and Adelson E.H. Laplacian Pyramid as a compact image code *IEEE Trans. on Communication*,Vol COM-31 No -4 Apr 1983.

[4] Chan Y.T. *Wavelets Basics Kluwer Academic Press* 1995

[5] Charles K Chui *An introduction to wavelets Academic Press ,Inc.*

[6] Coifman R.R. and M V. Wickerhauser Best adapted wavelet packet bases *Preprint Yale university* Feb. 1990.

[7] Coifman R.R. and M.V. Wickerhauser Entropy based algorithm for bestbasis selection *IEEE trans. on information theory* Vol 38 pp 713-718 Mar1992.

[8] Daubechies I. *Ten lectures on wavelets Society for industrial and applied mathematics* 1992

[9] Daubechies I. Orthonormal bases of compactly supported wavelets *Comm. Pure Appl. Math.* Vol 41 pp 909-996,1988

[10] Epstein B.R., Hingorani R., Shapiro J. M. and Cziler M. KLT-wavelet data compression for Landsat thermatic mapper images *Proc. Data Compression Conf.* Apr. 1992, pp. 200-208

[11] Gupta S. and Gersho A. Feature predicted vector quantization of multispctral images. *IEEE trans on geoscience and remote sensing* Vol 30 No 3,May 1992

[12] Gonzales R C. and Wintz P. Digital Image Processing *Addison-Wisley publishing corp.* 1987

[13] Jain A K. Digital image processing

[14] Jain A. K. Image Data Compression : A review *Proc. IEEE* Vol 69, No. 3 Mar 1981 pp 349-389.

[15] Lillesand and Keifer *Remote sensing and image interpretation Wiley eastern & sons Inc.*

[16] Lim J. S. *2 dimensional signal processing*

[17] Mallet S. Theory for multiresolution signal decomposition: A wavelet representation newblock *IEEE trans. on pattern anal. & mach. Intell.*Vol.11 ,pp 647-693. july 1989

[18] Mallet S. Multifrequency channel decomposition of images and wavelet model *IEEE trans.ASSP* Vol.37, pp 2091 -2110,Dec 1990

[19] Netravali A.N. and Haskell B. G. Digital Pictures representation and compression *NewYork Plenum Press* 1988

[20] Netravali A N. and Limb J O. Picture coding :A review *Proc. of IEEE* mar 1980

[21] Rioul O and Vatterli M Wavelets and signal processing *IEEE signal processing mag.* Vol 8,pp 14-38 Oct 1991

[22] Rosenfeld A.and Kak A.C. Digital Picture Processing *New york: Academic Press* 1976

[23] *special issue of IEEE trans. on information theory* Mar 1992

[24] Saghari J. A., Tescher A.G and Reagan J. T. Practical Transform coding of multispectral imagery *IEEE signal processing magazine* Jan 1995.

[25] Shapiro J.M. Embedded image codoing using zerotrees of wavelet coefficients *IEEE trans. on signal processing* No.12 Dec 1993

[26] Vatterli M Wavelets and subband coding

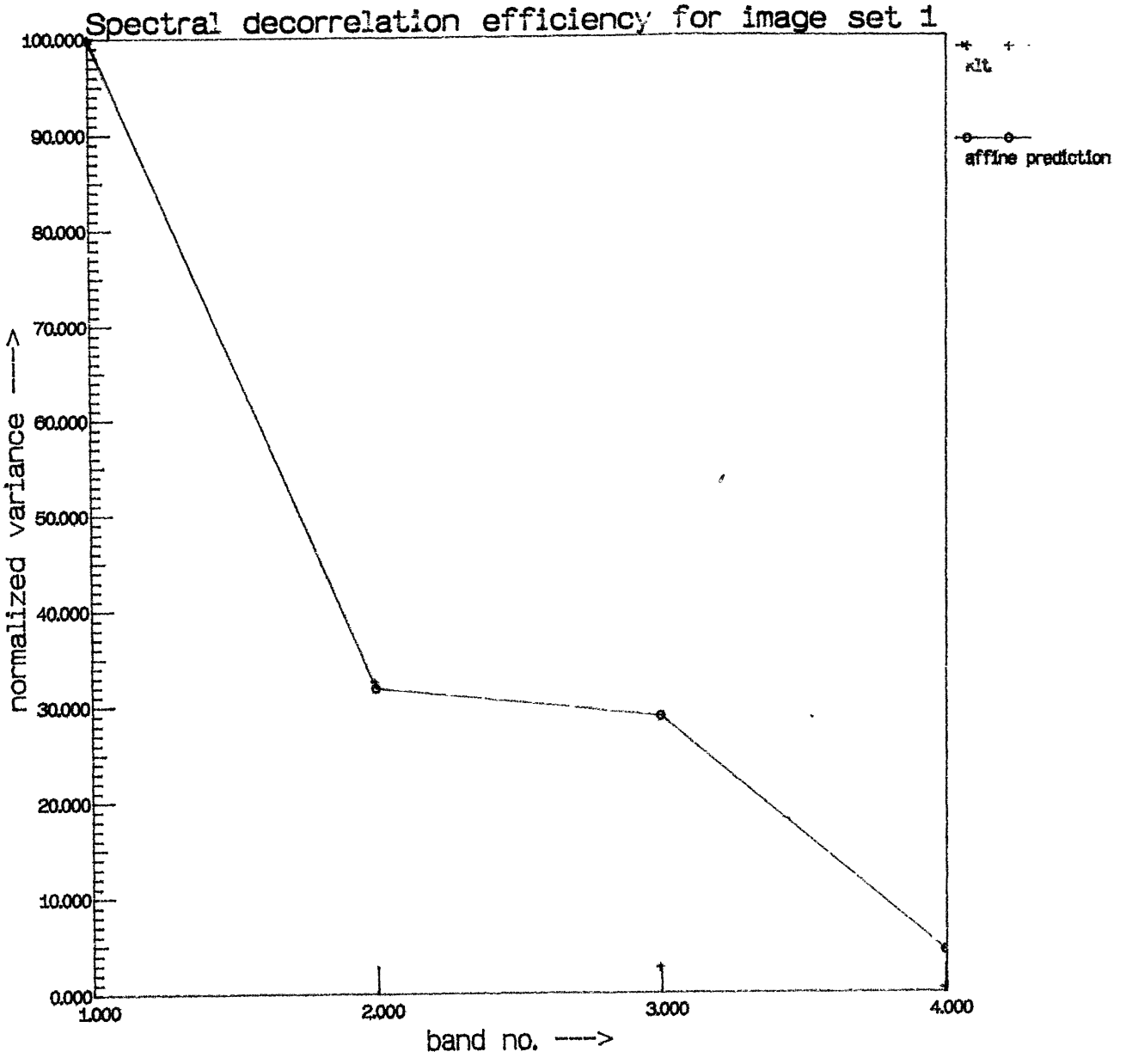
[27] Vatterli M. and Herley C. Wavelets and filterbanks *IEEE trans. on signal processing* Sep 1992

[28] Vaughen V.D. and Wilkinson T. S. System consideration for multispectral image compression *IEEE signal processing mag.* Jan 1995

[29] Witten I.H. ,Neal R. and Cleary J.G. Arithmetic coding for data compression *A.C.M. Vol30 ,pp 520-540* June1987

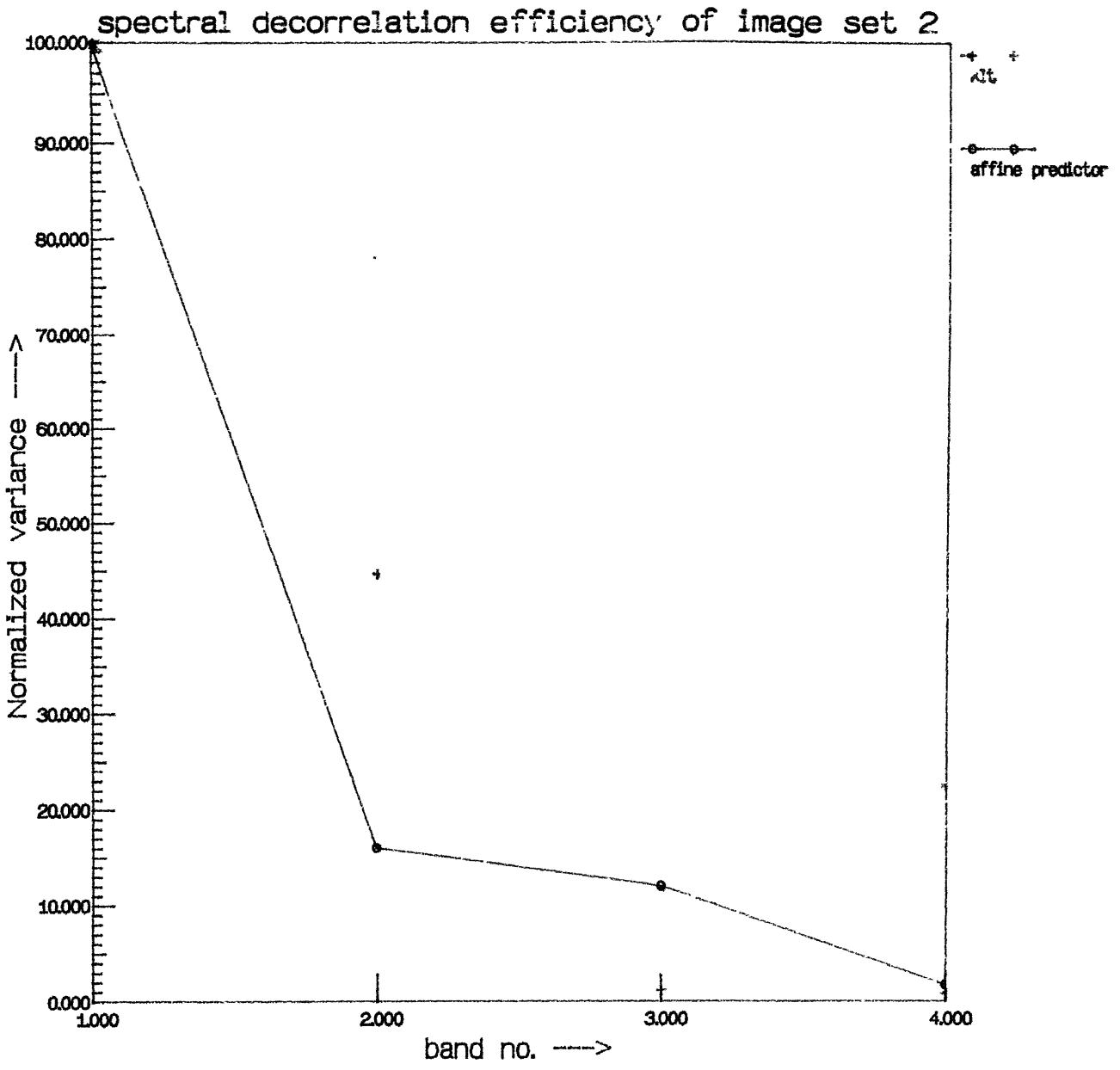
[30] Young R.K. Wavelet theory and its application *Kluwer Academic Press*

PLOT 1



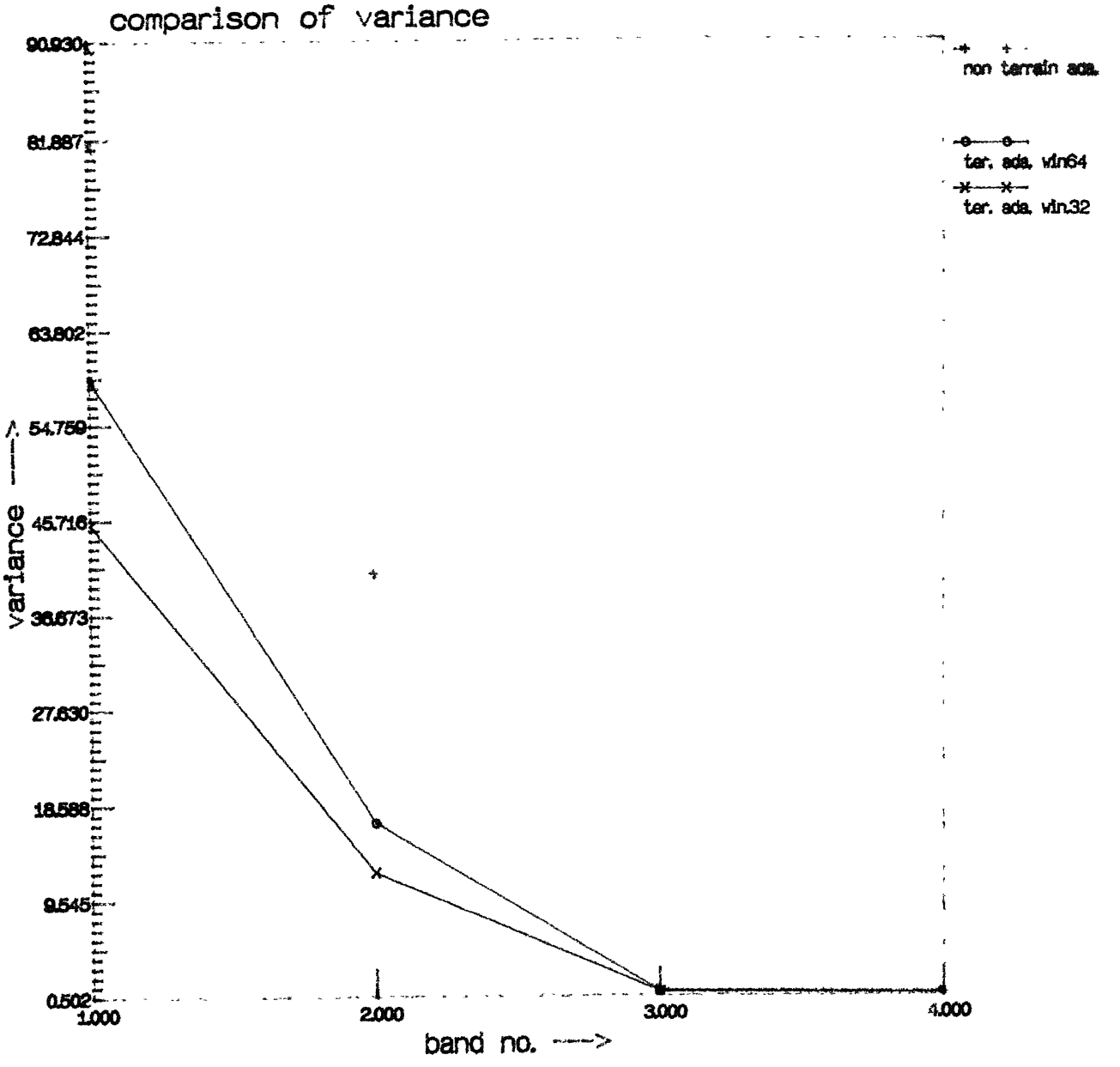
UsPlot

PLOT 2



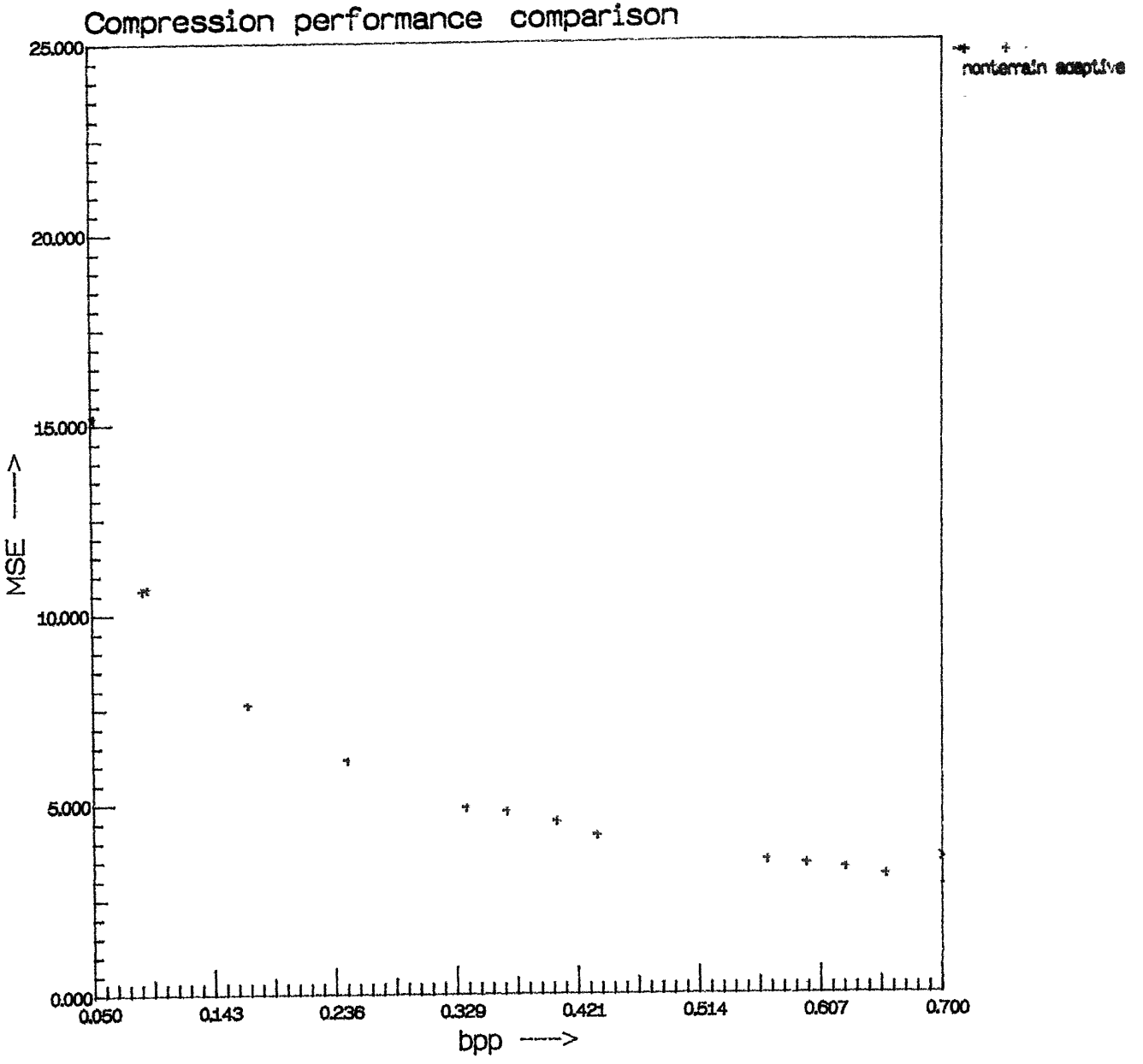
UsPlot

PLOT 3

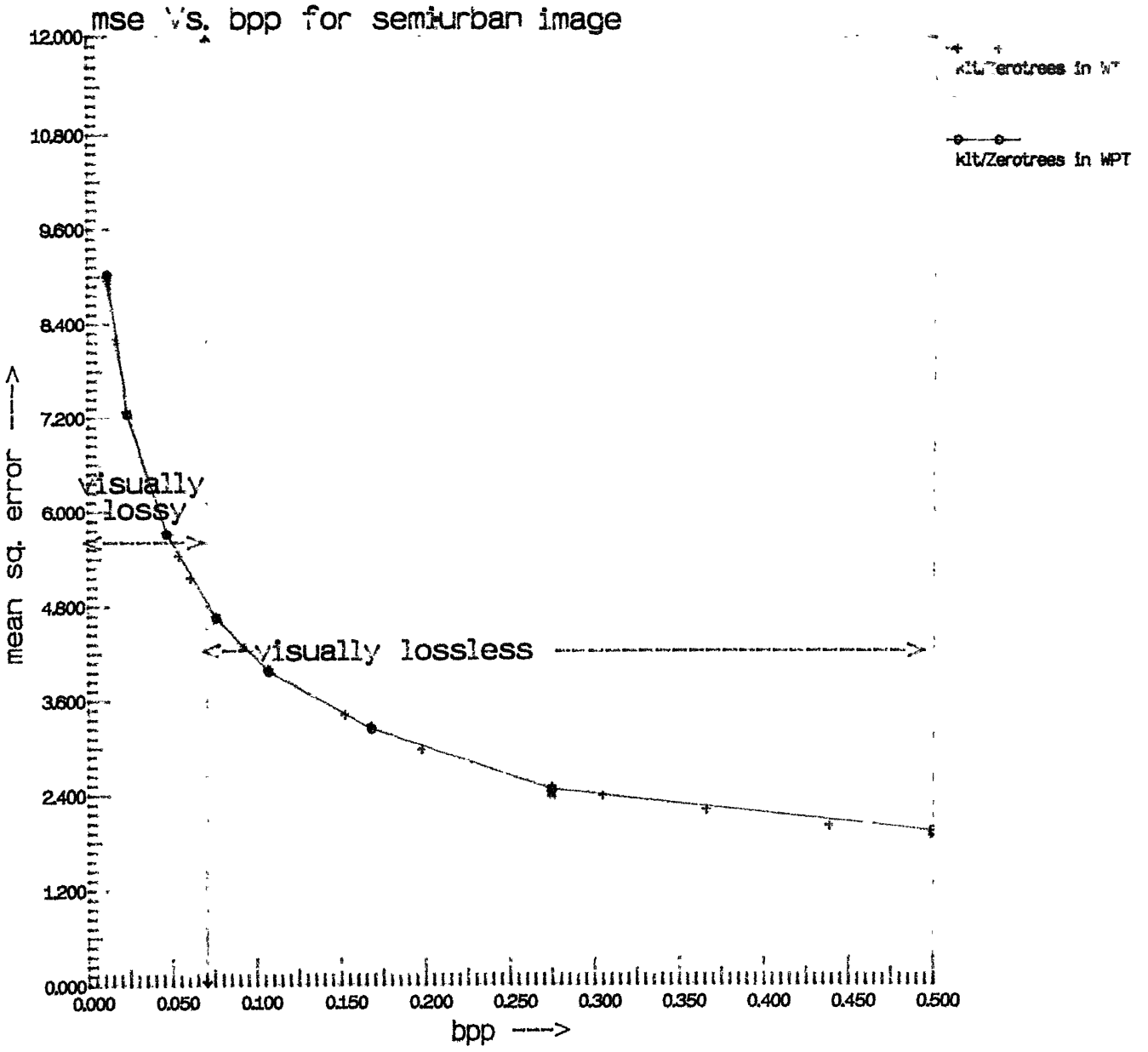


UsPlot

PLOT 4

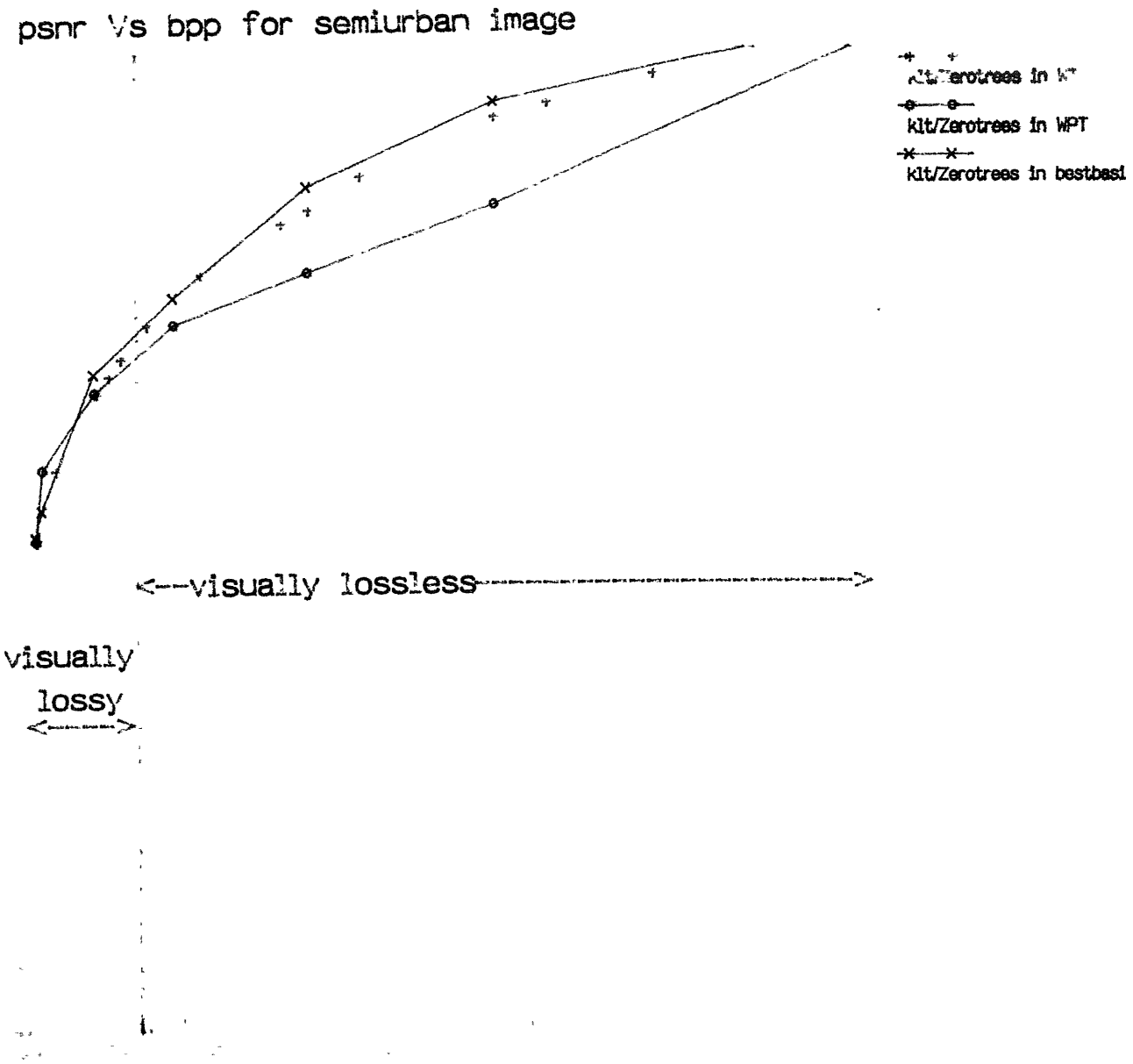

Compression performance comparison
nonterrain adaptive
25.000
20.000
15.000
10.000
5.000
0.000
MSE ---->
0.050
0.143
0.236
0.329
0.421
0.514
0.607
0.700
bpp ---->
UsPlot

PLOT 5



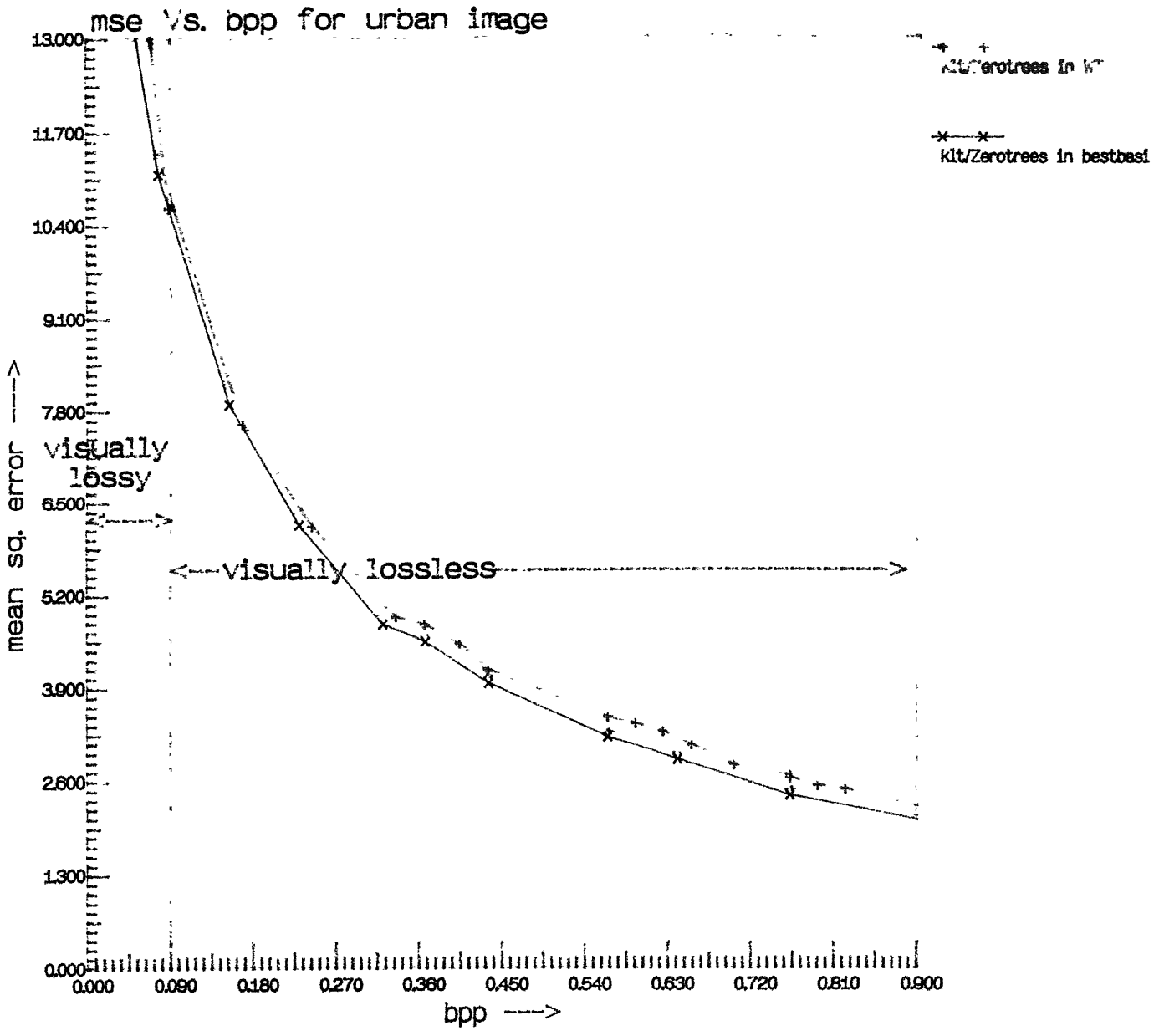
UsPlot

PLOT 6



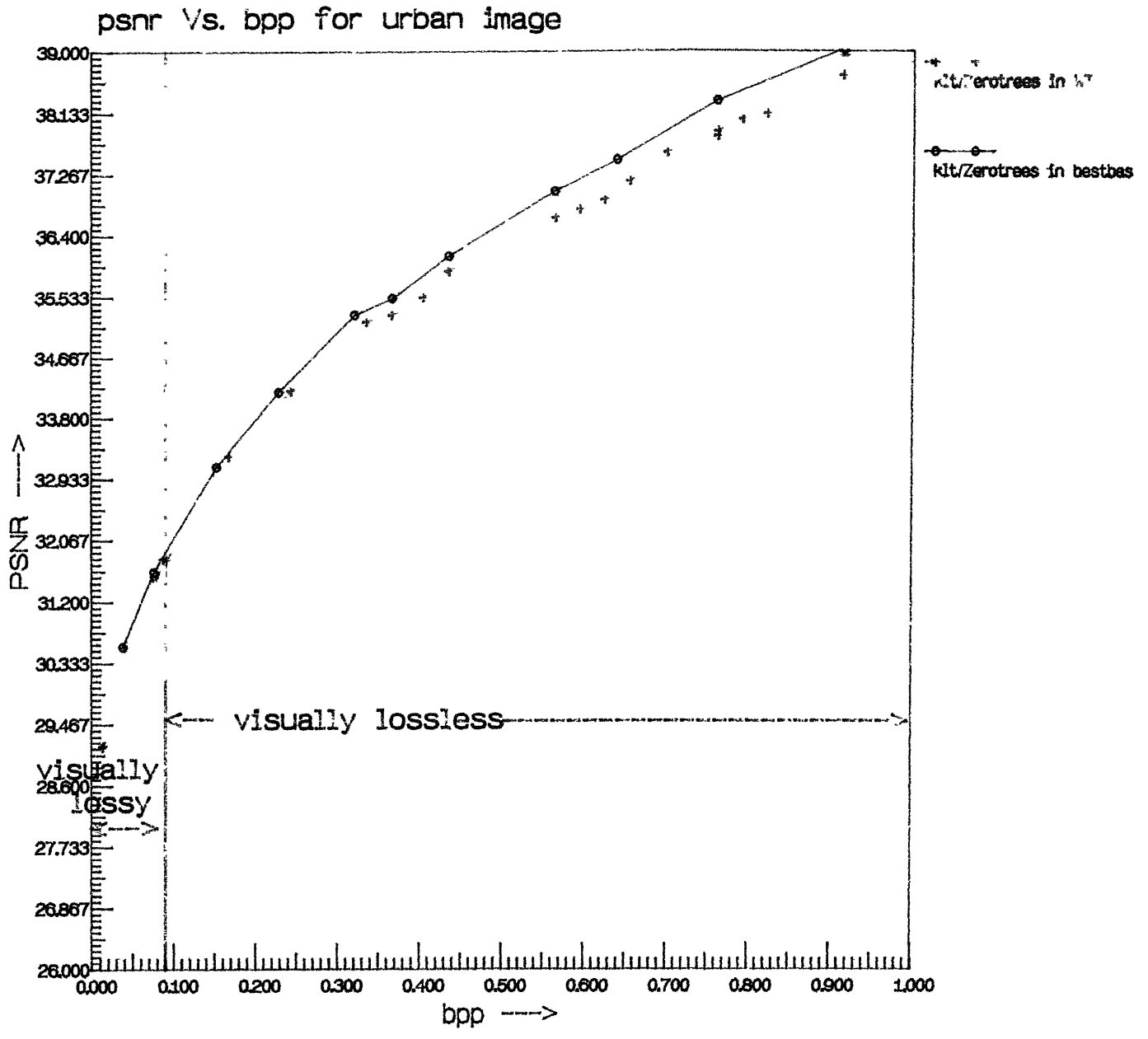
UsPlot

PLOT 7



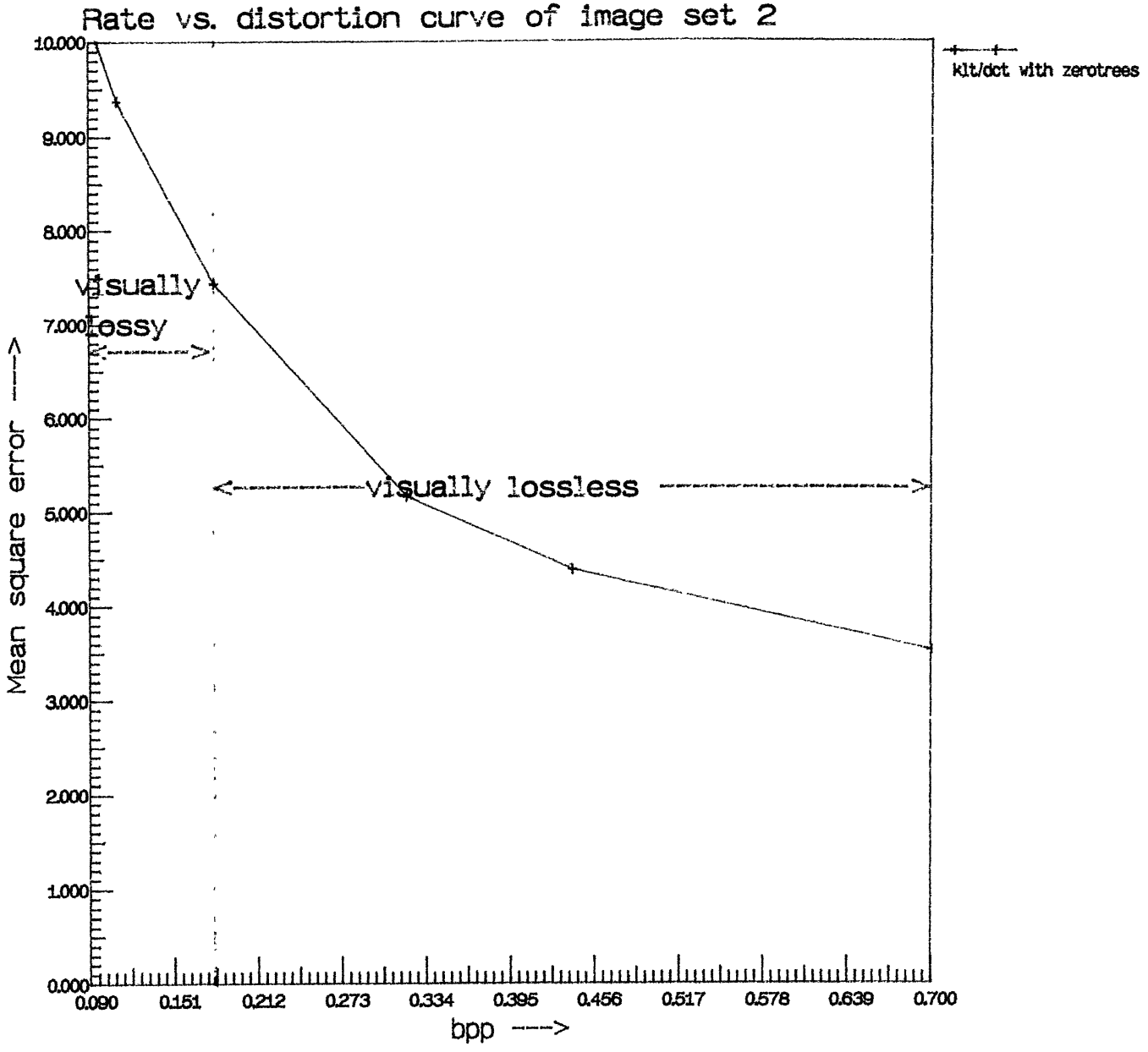
UsPlot

PLOT 8



UsPlot

PLOT 9



UsPlot